Applied Physics Laboratory
University of Washington

FINAL TECHNICAL REPORT

ONR Grant N00014-15-1-2549
Critical Transitions and Adaptation in Group Dynamics

<u>Principal Investigator</u>
Dr. Michael Gabbay
Applied Physics Laboratory
University of Washington
1013 NE 40th Street
Seattle, WA 98105

<u>Period of Performance</u>: May 1, 2015 – September 30, 2018

Table of Contents

1 Scope ... 3
2 Accomplishments ... 3
3 Summary of Research ... 3
3.1 Frame-Induced Group Polarization Theory ... 4
3.2 ASC Model of Opinion Network Dynamics ... 6
3.3 Model Testing on Experimental Data ... 9
3.4 Integrated Experimental and Computational Approach toward Social Influence ... 11
3.5 Ideology and Militant Group Interactions ... 12
3.6 Terrorist Decision Making in Northern Ireland ... 13
4 Transition Activity ... 15
5 Training ... 15
6 Publications and Presentations ... 15
6.1 Publications ... 15
6.2 Presentations ... 16

Appendices

Appendix 1: Opinion Network Modeling and Experiment

Appendix 2: Networks of Cooperation: Rebel Alliances in Fragmented Civil Wars

Appendix 3: Integrating Computational Modeling and Experiments: Toward a More Unified Theory of Social Influence

Appendix 4: Fratricide in Rebel Movements: A Network Analysis of Syrian Militant Infighting

1 Scope

This report describes research conducted by the Applied Physics Laboratory of the University of Washington (APL-UW) for the Office of Naval Research (ONR) under grant N00014-15-1-2549. The research performed under this grant explored both the basic behavior of group opinion dynamics and aspects of militant group decision making and interactions. Opinion models were developed, simulated, and tested against experimental results. Empirical data on militant groups in Syria and Northern Ireland were collected and analyzed. The principal investigator (PI) of this effort research was Dr. Michael Gabbay.

2 Accomplishments

The following are significant accomplishments of this effort:

1. Developed a novel theory of frame-induced group polarization that is grounded in the theory of decision making under risk and uncertainty. (Sec. 3.1)
2. Developed a novel model of opinion network dynamics, the accept-shift-constrict (ASC) model, which enables the emergence and persistence of majorities comprised of proximate positions and makes a distinction between opinion and rhetorical positions. (Sec. 3.2)
3. Demonstrated that the ASC model can account for discussion-induced shifts to the extreme without the typical opinion network modeling assumption of greater resistance to persuasion among extremists. (Sec. 3.2)
4. Showed that ASC model simulations of large networks display a sharp transition between convergence to a common opinion and divergence into opposing camps. (Sec. 3.2)
5. Demonstrated the quantitative agreement of the ASC model (and the related but simpler RPM model) with the results of a group discussion experiment – a rare achievement in group dynamics research. (Sec. 3.3)
6. Demonstrated the ability of the RPM model to predict polarization out-of-sample. (Sec. 3.3)
7. Developed a proposed framework for the integration of computational and experimental approaches to social influence. (Sec. 3.4)
8. Developed a theoretical framework for the role of ideology in militant decision making with respect to cooperation and conflict with other groups. (Sec. 3.5)
9. Empirically demonstrated that Syrian militant networks are strongly shaped by ideology in accord with our theory using data collected by automated and manual content analysis. (Sec. 3.5)
10. Collected data on decision making by militant groups in Northern Ireland and identified factors responsible for escalation or de-escalation decisions. (Sec. 3.6)

3 Summary of Research

In this section, we summarize key elements of this research effort. Reference numbers refer to documents listed in Section 6.1.

3.1 Frame-Induced Group Polarization Theory

Building upon an experiment and initial development done under the preceding grant (HDTRA1-10-1-0075, jointly funded by ONR), we developed a novel *frame-induced* theory of group polarization grounded in research on social influence and the theory of decision making under risk and uncertainty [5]. In the group polarization effect, also known as the risky shift, discussion among group members who are all on the same side of an issue induces more extreme decisions or opinions. Our frame-induced polarization theory is complementary to the two standard social psychology theories of group polarization: informational influence, in which discussion exposes like-minded individuals to new information in support of their side, and normative influence, in which a culturally-salient norm, such as one favoring risk taking, exerts peer pressure pushing groups toward the extreme. However, unlike those theories, frame-induced polarization is integrated with the concurrent social influence processes of majority influence and consensus pressure. Consequently, it can make predictions as to whether polarization will occur for groups with specified initial opinions whereas the standard theories always predict polarization whenever the polarization preconditions are present, regardless of the distribution of initial opinions within the group. This uniform polarization prediction is also a problem of the "extremist tilting" mechanism employed in the opinion network modeling literature in which individuals with more extreme positions are taken to be more confident and hence exert outsized influence on the group. In addition, the frame-induced theory takes better account of the reference point, which defines the boundary between opposing issue sides, and communication network structure than existing theories. These properties enable the frame-induced theory to explain the results of our experiment on the effects of group discussion on football wagering, which are at odds with the informational influence, normative influence, and extremist-tilting theories.

In frame-induced polarization, the distinction between the quantitative policy under debate and the *rhetorical frame* – the aspect of the policy upon which deliberations focus – is central. The rhetorical frame will typically correspond to the dominant source of disagreement within the group due, for instance, to uncertainty as to the likelihood of an outcome. The rhetorical frame position $\rho(x)$ is taken to be a function of the policy x. Groups will tend to shift toward the extreme if the functional relationship between the rhetorical position and the policy is concave (negative curvature), i.e., the rhetorical position increases more slowly as the policy becomes more extreme.

A concave rhetorical function can yield polarization via *distribution reshaping*. The effect of concavity is to compress rhetorical distances toward the extreme relative to the distances between more moderate members, making it easier for majorities to form on the extreme side of the mean. Consequently, while the policy distribution may be symmetric so that no majority is favored on either side of the mean, the distribution of rhetorical positions is skewed so that there is an initial majority on the extreme side of the rhetorical mean. This *rhetorically-proximate majority* (RPM) converges to a policy position more extreme than the mean to which the moderate minority of group members (those with policies below the mean) then concurs, thereby resulting in a consensus policy that exhibits group polarization toward the extreme.

The members of the F group in Figure 1 show how polarization can occur by the combination of distribution reshaping and the RPM process. The positive and negative sides of the horizontal axis signify opposing pro and con sides of the policy respectively. The F group members are located on the positive policy axis indicating that they are homogeneous in terms of which policy side they prefer. The rhetorical function is concave in that as the policy becomes more extreme the incremental change in the rhetorical position diminishes (stated this way, the rhetorical function can be seen to be concave for both policy sides – the overall S-shape is an artifact of using negative values to represent the con side). Although the intermediate member F_2 is equidistant in policy from the moderate F_1 and the extremist F_3, F_2 is rhetorically closer to F_3 and therefore (F_2, F_3) is the RPM pair. They agree on a policy halfway between them to which F_1 comes up due to majority influence. The RPM policy is seen to be greater than the initial mean policy.

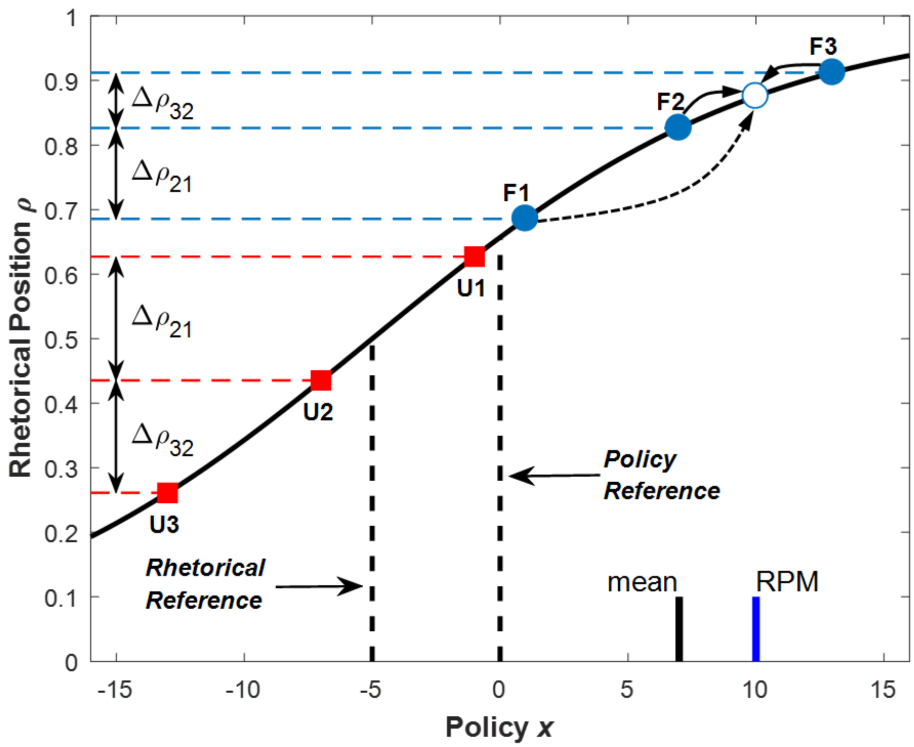

Figure 1. Illustration of distribution reshaping, RPM process, and reference point shifting due to heuristic frame substitution. Rhetorical function relating rhetorical frame position to policy is solid curve. Solid curved arrows indicate formation of the RPM pair by F_2 and F_3; dashed curved arrow indicates conformity of F_1 to form final consensus. Short black line at bottom is the mean of the initial F group policies; short blue line is RPM process prediction for the F group consensus policy.

We justified the expectation of a concave rhetorical function using the theory of decision making under risk and uncertainty. In particular, the rhetorical function was derived for the case where the policy is a continuous variable (e.g., a wager) which depends on the forecast of a binary event (a game outcome). For a binary gamble in which the quantitative policy depends on the likelihood of one of the two outcomes, the rhetorical frame should be the subjective probability that that outcome will occur. Using the standard assumption in the decision making literature that most individuals are risk averse, this subjective probability can be shown to be a concave function of the policy. Therefore, polarization via distribution reshaping is expected.

In addition to distribution reshaping, the frame-induced theory also incorporates *heuristic frame substitution*, a behavior that can account for unequal polarization by policy side so that, for instance, one policy side may show systematic polarization whereas the other does not (as observed in the experiment). In heuristic frame substitution, a simpler, more intuitive rhetorical frame is discussed in place of a more complex frame that directly corresponds to the policy. Heuristic frame substitution can shift the reference point of the rhetorical frame away from the policy reference as indicated in Figure 1. The rhetorical reference splits the con (negative) policy side, which results in the U group being arrayed on the approximately linear part of the rhetorical function rather than on the shoulder as for the F group. Consequently, U_2 is roughly the same rhetorical distance from both U_1 and U_3. Considering the effects of uncertainty and noise, formation of the moderate (U_1, U_2) RPM pair is about as likely as the extreme (U_2, U_3) pair so that systematic group polarization is absent or much reduced.

Our integration of decision making under risk and uncertainty into the frame-induced polarization theory allows us to provide specific guidance as to when heuristic frame substitution is expected. In the binary gamble context, such substitution can occur when there are two distinct gambles that depend on the same random variable but with different thresholds: the policy gamble that directly determines whether one's policy choice is successful and a heuristic gamble that is more intuitively accessible.

3.2 ASC Model of Opinion Network Dynamics

We developed two mathematical models consistent with frame-induced polarization [1, 5]. The RPM process described above is formulated as the RPM model, which uses a simple aggregation procedure to determine the group consensus policy via a weighted average of the policies of the majority of group members whose rhetorical positions span the least range [5]. Network structure is accommodated by weighting policies by relative node degrees. We also developed a new model of group opinion dynamics on networks which can account for experimentally-observed shifts toward the extreme better than existing models (see Sec. 3.3). The *accept-shift-constrict* (ASC) model describes the opinion change process on a network over time as a result of dyadic-level interactions [1]. The ASC model makes two innovations beyond existing continuous opinion network models. First, it makes a distinction between opinion (or policy, more generally) and rhetoric in accord with the frame-induced theory of group polarization. Second, it incorporates a novel uncertainty reduction mechanism which does not require that node uncertainties be visible to others. As a consequence, the ASC model can produce discussion-induced shifts toward the extreme without reliance upon systematic skews in individual psychological traits such as stubbornness as has been typically assumed within the opinion network modeling literature on extremism.

The ASC model assumes an underlying dyadic process in which one node sends a message to a receiver node in an effort to persuade the latter. The message can impact both the receiver node's policy and its uncertainty interval around that policy. Conceptually, the model proceeds in distinct accept, shift, and constrict phases (although all occur simultaneously in the mathematical formulation). The accept and shift phases occur in the equation that governs the rate of change of the node's policy. In the accept phase, the ASC model assumes that the probability that the receiver node will accept the message as persuasive decreases as a Gaussian function of the *rhetorical* distance between the sender and receiver nodes. The uncertainty of the receiver's position is taken to be the standard deviation parameter in the Gaussian. If a message is accepted, then, in the shift phase, the receiver shifts its policy in the direction of the sender's by an amount proportional to their *policy* difference. The constrict phase is governed by a second equation for the rate of change of a node's uncertainty, modeling a process in which interaction with others with close positions reduces uncertainty. If the sender's rhetorical position is within the uncertainty interval of the receiver, then the receiver decreases its uncertainty but not below a certain minimum value. Accordingly, unlike other models that involve uncertainty dynamics, it is the difference in (rhetorical) positions among dyad members rather than their difference in uncertainties that drives uncertainty change. The network weights in the ASC model represent the influence of one node upon another due to factors such as communication rate and expertise; they need not be symmetric. The ASC model is implemented in terms of coupled nonlinear ordinary differential equations, with two equations for each group member, one for the policy and one for the uncertainty.

A crucial consequence of the uncertainty reduction dynamics in the ASC model is the ability for interim majorities to more effectively maintain their position in the face of minority influence. A sample simulation of the ASC model as applied to a triad is shown in Figure 2a. All the nodes are connected and the network weights between nodes are symmetric and all equal. The initial policies are set such that the middle node policy, x_2 , is closer to x_3 than x_1. We observe that x_2 and x_3 form a majority and their uncertainties (λ_2, λ_3) quickly reach their minimum values while that of the minority member (λ_1) stays at its initial value. Consequently, x_1 is more open to accepting messages from the majority pair than vice versa, so that x_1 essentially comes up to the majority position, resulting in a consensus policy that is shifted upward from the mean. This ability for a majority to emerge and persist in the face of minority influence is critical to the frame-induced mechanism of group polarization. This dynamic, however, is not present in linear opinion network models such as the DeGroot, Friedkin-Johnsen, and consensus protocol models. As shown in Figure 2b for the consensus protocol (also known as the Abelson model, a continuous time equivalent of the DeGroot model), no interim majority emerges as all nodes converge simultaneously on the initial mean.

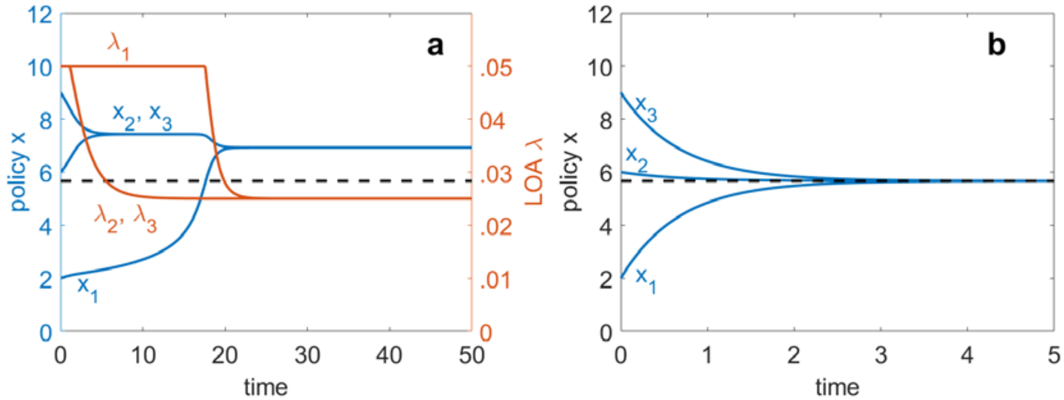

Figure 2. Evolution of policy positions and uncertainties for completely-connected triad. (a) ASC model. (b) Consensus protocol. Dashed lines indicate initial mean policy.

We also showed that the ASC model could generate critical transitions between population convergence and divergence for large networks. The control parameter that induces the transition is the risk aversion. For low risk aversion values, an initially unimodal (Gaussian) opinion distribution converges to a common opinion at the center of the distribution (Figure 3a). For high risk aversion values, the same initial opinion distribution results in a divergence into two opposing camps (Figure 3b). The simulation was run until member opinions remained level for a substantial time period (eventually, if the model is run long enough, consensus will always be reached). As the risk aversion is varied, the convergent behavior persists until a critical value, beyond which the opinion discord between the opposing camps rises rapidly (Figure 3c). The implication of this result is that allowing discussion between all the members in a large group can actually increase dissension within the group (for sufficiently high risk aversion). This work has not yet been published.

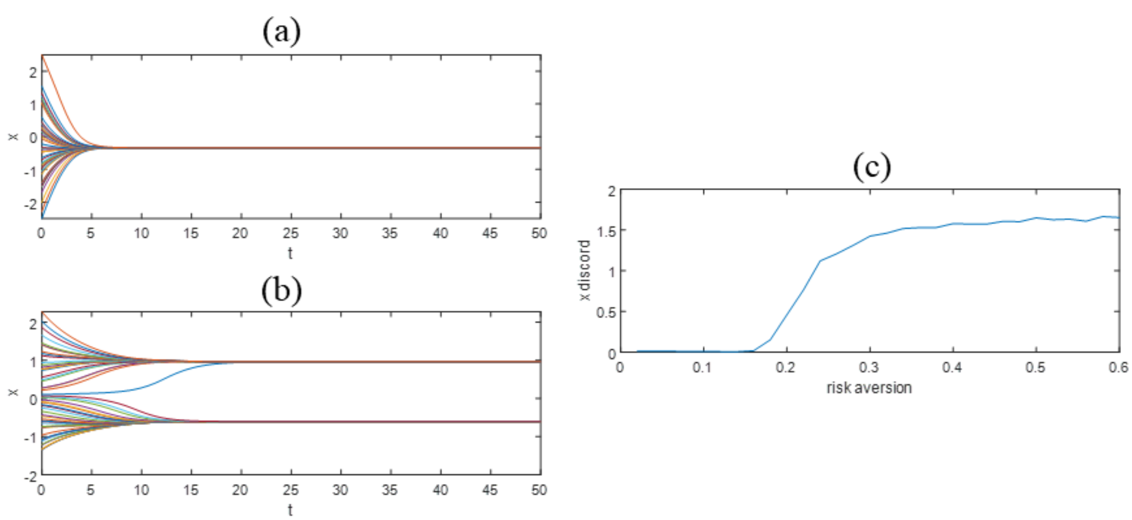

Figure 3. ASC model simulations for large network showing transition between convergent and divergent behavior. Opinion vs. time for (a) low risk aversion of 0.1 and (b) high risk aversion of 0.35. (c) Final policy discord vs. risk aversion showing transition between convergence and divergence at risk aversion of 0.16. A complete network topology of 51 nodes was used.

3.3 Model Testing on Experimental Data

We briefly describe the experiment presented in [5] used to test our theory and models. Three-person groups engaged in online discussion about wagering on National Football League (NFL) games. As is standard practice in football betting, spread betting was employed rather than wagering directly on which team will win the game. In spread betting, the terms "favorite" and "underdog" are used to refer, respectively, to the likely winner and loser of the game itself. The point spread is the expected margin of victory of the favorite team as set by Las Vegas oddsmakers. A bet on the favorite is successful if its margin of victory exceeds the spread; otherwise a bet on the underdog is successful. On the basis of a pre-survey that elicited team choices and wager amounts, discussion groups were constructed with respect to three dichotomous variables: (1) *policy side* of favorite or underdog corresponding to the team chosen as more likely to beat the spread; (2) *disagreement level* of "high" or "low" between the minimum and maximum wagers in the group; and (3) *network structure* of "complete" in which all members could communicate with each other or "chain" in which the members with the lowest and highest wagers were the end nodes of the chain and the intermediate-wager member served as the center node. After discussion, each member made their final wager. A group decision was not required but groups arrived at a consensus wager far more often than the alternative outcomes of a two-person majority or three different wagers.

Defining an increase in mean wager due to discussion as group polarization (risky shift), statistically significant results were observed in consensus groups for all three variables: (1) systematic polarization only on one policy side (favorite), and greater polarization for (2) high disagreement groups (vs. low) and (3) complete networks (vs. chains). While these results are not readily explained by existing group polarization theory, they are in accord with our frame-induced theory and the qualitative behavior exhibited in RPM and ASC model simulations. The first result is consistent with heuristic frame substitution, the second with distribution reshaping and rhetorically-proximate majority formation, and the third with the greater centrality and intermediate initial wager of the chain middle node.

Furthermore, we showed that the RPM and ASC models are in quantitative agreement with the experimental results. The models were simulated using actual group initial wagers to predict consensus wagers. Free parameters were fit to minimize the error with the data. The mean wager shift due to group discussion as a function of the difference in wagers for the experimental data and simulations is plotted in Figure 4. Both models were found to pass a statistical goodness-of-fit test. Alternative models (the median, policy-based proximate majority, policy averaging weighted by rhetorically-based confidence, and extremist-tilting versions of the Abelson model, one weighted by wager confidence and the other by rhetorically-based confidence) did not pass the goodness-of-fit test [1,5,7].

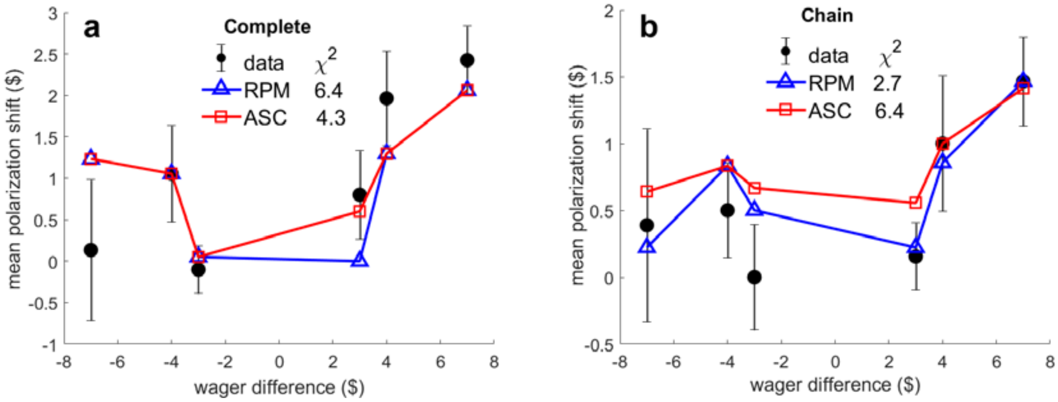

Figure 4. Comparison of experimental data and simulations of ASC and RPM models. (a) Complete network. (b) Chain. Positive and negative sides of the x-axis correspond to favorite and underdog groups respectively. Error bars are standard errors. Simulation results are rounded upward to nearest dollar.

We also demonstrated the ability of the RPM model to successfully predict polarization out-of-sample (not published). A cross-validation procedure was employed in which a randomly chosen subset groups was used as the training set to fit the risk aversion parameter (the only free parameter in the RPM model). The remaining groups were then used as a test set to evaluate if the model prediction agrees with the data as to whether the group post-discussion consensus wager will be greater than the pre-discussion mean. The histogram of RPM model accuracy over 500 repetitions is shown in Figure 5. The RPM model was found to have a mean accuracy rate of 0.69 with a 95% confidence interval of (0.56, 0.83) indicating statistically greater than chance performance. Alternative models, such as the median and wager-based proximate majority, did not yield accuracy statistically greater than chance.

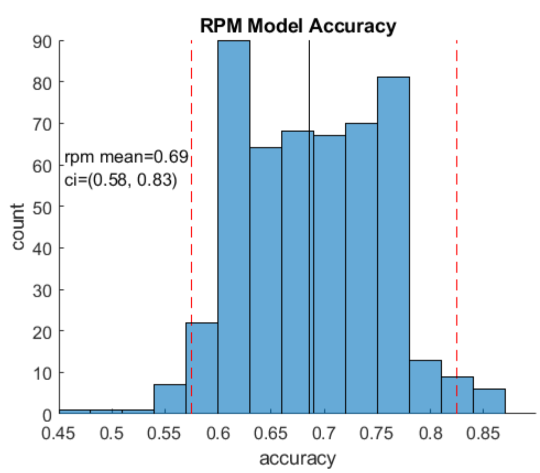

Figure 5. Distribution of RPM model accuracy in predicting polarization out of sample over 500 samples. Training set size = 40 groups. Test set = 119 groups.

3.4 Integrated Experimental and Computational Approach toward Social Influence

We proposed a research framework which calls for more dedicated integration of computational modeling and experimental approaches toward studying social influence phenomena [3]. This framework is intended to build upon and replicate the relatively rare achievement described above of demonstrating the quantitative agreement of theoretically-driven mathematical models with experimental data. This work was done in conjunction with a DARPA-sponsored workshop on the Future of Computational Social Science and the associated edited volume.

An overview of the framework is shown in Figure 6. Its goal is the quantitative testing of computational (not statistical) models of social influence. Of particular importance to this approach are complex systems models, such as the ASC model, in which the variables of interest (e.g. group member opinions) evolve from initial conditions as a result of endogenous feedback with each other and perhaps exogenous signals. The approach centers upon the conduct of experiments explicitly designed to test the quantitative predictions of models rather than the standard experimental paradigm of testing qualitative hypotheses. Its aim is the development of broad models that can account for a range of phenomena and experimental results. Elements of this approach include: exercising discipline and discrimination with respect to model parameters; conducting goodness-of-fit tests; more highly resolved initial variable conditions; more deliberate control of initial opinion distributions; measuring opinions or other variables over time; greater use of out-of-sample prediction; testing models on new and old data to foster model convergence not proliferation; and parameterizing the nature of group tasks along a spectrum rather than ambiguously assigning them to nominal categories such as intellective or judgmental.

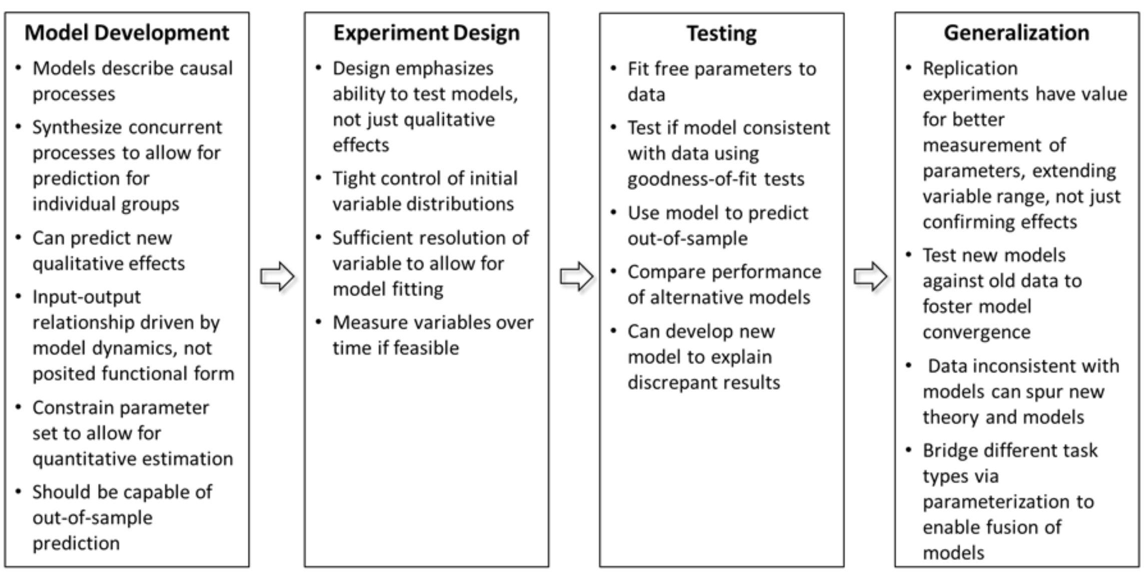

Figure 6. Overview of integrated modeling-experiment approach.

A major advantage of this integrated approach is an improved ability to synthesize different effects. Models can synthesize multiple effects more readily than combining different, often ambiguous, categorizations of conditions. Other anticipated benefits include: raising the bar for the evaluation of rival theories with higher precedence given to theories whose associated models are in quantitative accord with experiment; stronger incentive to conduct experiments for the purpose of providing better measurement or expanding the range of model variables, with ensuing greater replicability; greater ability to publish anomalous findings thereby spurring new theory and model development; and, on an applications level, raising the confidence and scope with which models can be applied to natural situations for purposes of both prediction and designing interventions to shape outcomes.

3.5 Ideology and Militant Group Interactions

We developed and empirically tested a theory for how ideology influences militant decision making with respect to cooperation and conflict with other groups [2, 4]. Initial theory development and some data collection and analysis were done under this effort. Conduct of this research was transferred to ONR grant N00014-16-1-2919 upon commencement of that effort.

Central to the theory is the identification of three distinct components of militant ideology: the *conflict frame,* in which a rebel group identifies whom it is primarily fighting for and against, casting both ingroup and outgroups with respect to its preferred cleavage whether ethnic, religious, economic, or political; the *conception of the ideal polity,* in which a group identifies its vision for the post-conflict social and political order; and the *territorial aspiration,* in which a group identifies the geographical domain of this future order. We hypothesized that similarity along these three dimensions would facilitate cooperation among groups and dissimilarity would fan conflict. Our positing of an important role for ideology in shaping militant cooperation stands in contrast to prevailing realist views of militant alliance formation in which the distribution of power is afforded almost exclusive causal power.

We collected data on Syrian militant groups. We combined automated text processing and expert coding to generate quantitative data from several thousand statements of groups' activities and operational claims. These statements come from US Government translations of insurgents' statements and operational claims, drawn from social media (Facebook, Twitter and YouTube) as well as various jihadist forums and radio broadcasts. Automated methods were used in the collection of data on joint operations between militant groups, ideology, and claimed target types. Manual coding was used to generate scores along the above three ideology components. Reports by think tanks and other sources were used to assess group power (number of fighters) and state sponsorship. The full joint operations network consisted of 220 groups. However, the network analysis was conducted on a smaller set of groups for which attribute data (ideology, power, state sponsorship) was available. Figure 7 displays the joint operations network along with group ideological designations.

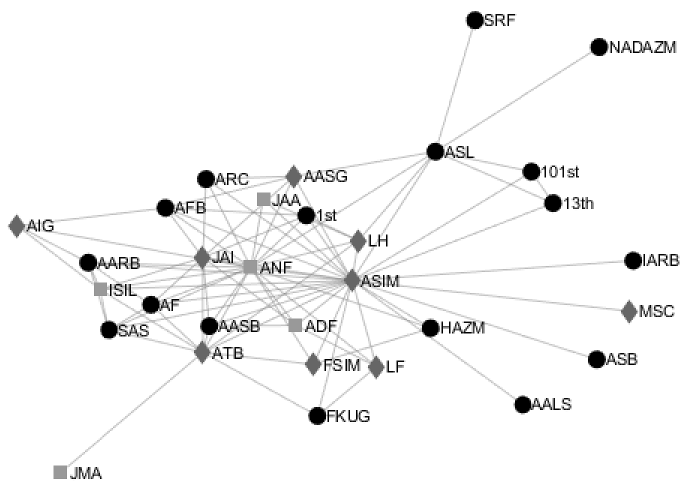

Figure 7. Diagram of joint operations among 31 Syrian militant groups. A line between two groups indicates the presence of at least one joint operation between them. Node shapes denote secular nationalist (circles), Salafist nationalist (diamonds), and sectarian jihadist (squares) ideological classifications.

A variety of network analysis methods and metrics including community structure, assortativity, network simulation, latent space modeling, and exponential random graph modeling were used to evaluate the relationship of network structure to ideology, power, and state sponsorship. Ideological similarity was found to be a highly significant driver of cooperative network structure in that groups who were ideologically closer tended to cooperate more frequently [2]. This held true for two of the three components (conflict frame and territorial aspiration) and the average of all three. No consistent support for how power is related to network structure was found. No statistically significant support for the proposition that having a common state sponsor would increase cooperation among groups was found. In accord with our theory that conflict frames help circumscribe the set of permissible targets, we also found that groups with similar conflict frames had more closely aligned targeting portfolios, i.e., the types of targets they claim such as state security forces, rival ethnic militias, government workers, and civilians (not published). A parallel analysis of the network of armed clashes between groups showed that ideological dissimilarity made such infighting more likely [4]. Larger power disparity between groups also significantly raised the level of infighting. Sharing a state sponsor, however, had no statistically significant effect on infighting.

3.6 Terrorist Decision Making in Northern Ireland

We developed a coding scheme for terrorist group decision antecedents and outcomes and constructed an associated dataset for Northern Ireland. The coding scheme is designed to help analyze the causes of major decisions by terrorist groups which represent substantial changes in policy or structure such as escalations in violence and group splits. It includes exogenous factors such as the level of repression or external shocks and internal factors such as changes in senior leadership and the power balance between hawks and doves. The scheme was implemented for groups on the republican and loyalist sides using narrative accounts of the conflict in Northern Ireland.

Using this dataset on decision factors and outcomes for militant groups in Northern Ireland, we conducted an analysis to identify the factors most responsible for decisions to escalate or de-escalate the level of violence. The method used was the information theory-based "reducing uncertainty model" of Drozdova and Gaubatz (International Studies Quarterly, 2014) designed to analyze small-N cases. This method quantifies the impact of various factors over a number of cases by calculating the mutual information of each factor and the outcome. Factors with high mutual information reduce the outcome uncertainty more than other factors and thereby are more important in predicting outcomes. This involves binarizing all the variables to analyze the impact each had on the outcome variable. For escalation, the biggest factor associated with escalatory decisions was an increase in resources. For de-escalation, the most important factor was that having hawks as the dominant leadership faction inhibited de-escalatory decisions. Interestingly, an increase in civilian support for a group did not appreciably affect its escalatory decisions but did appreciably act in opposition to de-escalatory decisions. For the Provisional Irish Republican Army, the main militant group, Figure 8 plots decisions to escalate or de-escalate the violence over time and its relationship to whether hawks or doves were in power. This work has not yet been published.

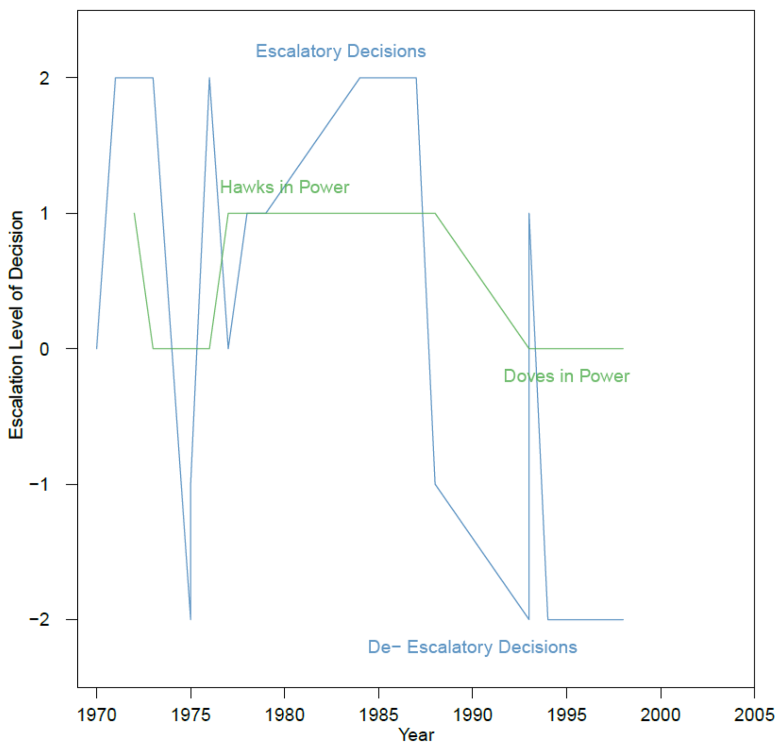

Figure 8. Escalation and de-escalation decisions and dominant leadership faction over time for the Provisional Irish Republican Army. The escalation level (blue curve) is on a five-point scale range from -2 (major de-escalation) to 2 (major escalation). The green curve shows the dominant faction and is equal to 1 for hawks in power and 0 for doves.

4 Transition Activity

The following is a list of the activity related to potential transition of this research:

1. PI held discussions with STRATCOM about incorporating elements of the research conducted under this project into the PORTEND software for the analysis and modeling of leadership networks and decision making [6].

2. In August 2016, PI briefed US government analysts on PORTEND methodology and participated in discussions about results of PORTEND application in collaboration with STRATCOM to two case studies.

3. In July 2017, PI briefed Defense Threat Reduction Agency (DTRA) personnel on this research.

4. In December 2018, PI briefed US government analysts and methodologists on this research.

5. In December 2018, PI met with Army Research Institute (ARI) personnel about this research and potential extensions.

5 Training

This grant partially supported the research assistantship of Emily K. Gade, a graduate student in the Department of Political Science at the University of Washington. She received her Ph.D. in June 2017.

6 Publications and Presentations

6.1 Publications

The accepted versions of forthcoming manuscripts are included in the Appendices as indicated. The following list omits several conference papers that are earlier versions of forthcoming papers.

1. Gabbay, M. (forthcoming). Opinion Network Modeling and Experiment. In V. In, P. Longhini, & A. Palacios (Eds.), *Proceedings of the 5th International Conference on Theory and Applications in Nonlinear Dynamics*: Springer. **(Appendix 1)**

2. Gade, E. K., Gabbay, M., Hafez, M. M., & Kelly, Z. (forthcoming). Networks of Cooperation: Rebel Alliances in Fragmented Civil Wars. *Journal of Conflict Resolution*. **(Appendix 2)**

3. Gabbay, M. (forthcoming). Integrating Computational Modeling and Experiments: Toward a More Unified Theory of Social Influence. In P. K. Davis, A. O'Mahony, & J. Pfautz (Eds.), *Social-Behavioral Modeling for Complex Systems*: Wiley. **(Appendix 3)**

4. Gade, E. K., Hafez, M. M., & Gabbay, M. (forthcoming). Fratricide in rebel movements: A network analysis of Syrian militant infighting. *Journal of Peace Research*. **(Appendix 4)**

5. Gabbay, M., Kelly, Z., Reedy, J., & Gastil, J. (2018). Frame-Induced Group Polarization in Small Discussion Networks. *Social Psychology Quarterly, 81*(3), 248-271.

6. Gabbay, M. (2018). Leadership Network Structure and Influence Dynamics. In E. Mitleton-Kelly, A. Paraskevas, & C. Day (Eds.), *Handbook of Research Methods in Complexity Science: Theory and Applications*. Cheltenham, UK: Edward Elgar Publishing.

7. Gabbay, M., Kelly, Z., Reedy, J., & Gastil, J. (2017). Rhetorically-Induced Group Polarization in Small Opinion Networks. *arXiv:1710.10295*.

6.2 Presentations

Presenter(s) listed in boldface.

1. **Gabbay, M**., "Opinion Network Modeling and Experiment," International Conference on Theory and Applications in Nonlinear Dynamics (ICAND), August 2018, Maui, HI.

2. **Gabbay, M**., "Integrating Experimental and Computational Approaches to Social Influence," Current Challenges in Computing workshop on Computational Social Science, September 2017, Napa, CA.

3. **Gabbay, M.**, Kelly, Z., Reedy, J., & Gastil, J., "Group Polarization due to Rhetorically-Induced Asymmetry and Heuristic Issue Substitution," poster presented at Political Psychology Section pre-conference of American Political Science Associations Annual Meeting, August 2017, Berkeley, CA.

4. **Gabbay, M.**, Kelly, Z., Reedy, J., & Gastil, J., "Group Polarization in Opinion Network Dynamics," North American Social Networks Conference, July 2017, Washington, DC.

5. **Gabbay, M**., "Extremism on Networks and Networks of Extremism," Defense Threat Reduction Agency, July 2017, Fort Belvoir, VA.

6. **Gabbay, M.**, Kelly, Z., Reedy, J., & Gastil, J., "Group Polarization in Opinion Network Dynamics," NetSci 2017, June 2017, Indianapolis, IN.

7. **Gabbay, M**., Kelly, Z., Reedy, J., & Gastil, J., "Choice Shift in Opinion Network Dynamics," American Physical Society March Meeting, Baltimore, MD, March 16, 2016.

8. **Zech, S., & Gabbay, M**., "Militant Networks as Political Networks," International Studies Association Annual Convention, Atlanta, GA, March 17, 2016.

9. **Gade, E.K**., Gabbay, M., Hafez, M., & Kelly, Z., "Bringing Ideology Back In: Militant Networks and Violence in Syria," International Studies Association Annual Convention, Atlanta, GA, March 17, 2016.

Appendix 1

Opinion Network Modeling and Experiment

Opinion Network Modeling and Experiment

Michael Gabbay

Abstract We present a model describing the temporal evolution of opinions due to interactions among a network of individuals. This Accept-Shift-Constrict (ASC) model is formulated in terms of coupled nonlinear differential equations for opinions and uncertainties. The ASC model dynamics allows for the emergence and persistence of majority positions so that the mean opinion can shift even for a symmetric network. The model also formulates a distinction between opinion and rhetoric in accordance with a recently proposed theory of the group polarization effect. This enables the modeling of discussion-induced shifts toward the extreme without the typical modeling assumption of greater resistance to persuasion among extremists. An experiment is described in which triads engaged in online discussion. Simulations show that the ASC model is in qualitative and quantitative agreement with the experimental data.

1 Introduction

While the experimental study of social influence and opinion change in particular primarily remains the province of the social sciences, the modeling of social influence dynamics, however, has extended into other fields including physics, computer science, and electrical engineering [1, 2, 3]. The primary goal of opinion network models is to predict final opinions from initial ones typically via a process that updates node opinions over time. Continuous opinion models — the concern of this paper — allow for incremental shifts in opinion where the amount of change depends upon the distance between node opinions and the network of interpersonal influence that couples nodes. The DeGroot and Friedkin-Johnsen models, as well as the consensus protocol (a continuous time version of the DeGroot model), use a

Michael Gabbay
Applied Physics Laboratory, University of Washington, 1013 NE 40th St, Seattle, WA, USA 98105-6698, e-mail: gabbay@uw.edu

linear dependence in which the shift is proportional to the opinion difference [4, 5, 6]. Bounded confidence models posit a hard opinion difference threshold, within which nodes interact linearly, but beyond which the interaction vanishes [7]. The nonlinear model of [8] uses a soft threshold so that, rather than vanishing completely, the interaction decays smoothly with distance.

Modeling how opinions become more extreme has been of particular concern in the opinion network modeling literature. The primary contribution of this paper is to present an opinion network model, the Accept-Shift-Constrict (ASC) model, which provides an experimentally supported depiction of the group polarization effect, a classic social psychology effect in which discussion among like-minded group members tends to make groups more extreme. The ASC model describes opinion change processes over a network as group members exchange messages. These processes consist of, first, the acceptance of a persuasive message which can then lead to a shift in the receiver's opinion and also a constriction of the receiver's uncertainty level. In turn, this constriction narrows the extent to which subsequent messages advocating distant opinions are accepted.

This paper proceeds as follows. The next section discusses the group polarization effect along with its treatment in social psychology and the opinion network modeling literature. Section 3 describes a recent experiment involving discussion about betting on National Football League (NFL) games, the results of which challenge existing group polarization theory. In Sec. 4, an alternative *frame-induced* theory of group polarization is presented that can account for the experimental results. Sections 5 and 6 present the ASC model and experimentally-relevant simulation results.

2 Group Polarization Effect

In the group polarization effect, discussion among group members who are all on the same side of an issue induces more extreme decisions or opinions ("polarization" as used here connotes a group shifting further toward one pole of an issue rather than diverging toward opposite poles as in conventional usage) [9, 10, 11]. It was originally referred to as the "risky shift" effect as it was discovered in an experimental context involving small groups faced with choosing among options of varying risk levels; discussion tended to shift groups toward riskier options than the average of their pre-discussion preferences. Subsequent research observed systematic discussion-induced extremism in homogeneous groups in broader contexts including social and political attitudes and the severity of punishments in jury deliberations. A group is considered to be homogeneous with respect to an issue if all its members have initial preferences that lie on one side of the issue's neutral reference point. Group polarization is then said to occur if after the discussion the mean preference of the group shifts further away from the reference point compared with the mean prior to discussion. Polarization is typically observed for issues that have a substantial judgmental component as opposed to issues like math problems that have demonstrably correct solutions.

Two distinct processes, based on informational and normative influence respectively, are most commonly accepted in social psychology as causes of group polarization [9, 10]. The informational influence explanation, known as persuasive arguments theory, focuses on the role of novel arguments. In essence, members of a homogeneous group, although inclined toward the same side of an issue, will typically possess different arguments in support of that side. The exchange of these arguments in discussion then exposes group members to even more information supporting their initial inclination and so shifts it further in the same direction. The normative influence explanation, social comparison theory, posits that the relationship of group member positions with respect to a culturally salient norm is critical rather than the information underlying those positions. The norm is taken to favor one pole of the issue. For example, a norm favoring risk-taking makes riskier positions more socially ideal than cautious ones. A major problem of the informational and normative influence theories is that they always predict polarization for an individual group whenever the polarization preconditions (homogeneous group and judgmental issue) are present, regardless of the distribution of initial opinions within the group. This problem stems from the fact that these theories were never reconciled with stronger, concurrent social influence phenomena such as majority influence and consensus pressure [12].

Within the opinion network modeling literature, extremism has been predominantly modeled by attributing higher network weights to nodes with more extreme initial opinions [13, 14]. This approach, which we refer to as “extremist-tilting,” is necessitated by the property of most continuous opinion models that the mean opinion in networks with symmetric coupling remains constant at its initial value — a property that is at odds with the shift in mean exhibited in group polarization. Consequently, extremists must be assigned greater influence over moderates than vice versa in order to shift the mean. This explanation is different from the two more prominent theories above but shares their problem of uniformly predicting polarization for homogeneous groups.

3 Experiment

This section describes the group polarization experiment conducted in [12] in which three-person groups engaged in online discussion about wagering on National Football League (NFL) games. As is standard practice in NFL betting, spread betting was employed rather than wagering directly on which team will win the game. In spread betting, the terms “favorite” and “underdog” refer, respectively, to the likely winner and loser of the game itself. The point spread is the expected margin of victory of the favorite team as set by Las Vegas oddsmakers. A bet on the favorite is successful if its margin of victory exceeds the spread; otherwise a bet on the underdog is successful.[1] If Team A is the favorite by a spread of six points over the underdog Team B, then

[1] In actual practice, bets are returned if the victory margin equals the spread.

Team A has to win the game by more than six points in order for a bet on Team A to pay off. The objective of the spread is to endeavor to equalize the odds for either the favorite or underdog to win the bet.

In the experiment, an upcoming NFL game was chosen and a pre-survey then elicited subject initial preferences with respect to team choice and a wager amount on that team from \$0 to \$7 (in whole-dollar increments). On the basis of the pre-survey, discussion groups were constructed with respect to three dichotomous variables. The first is *policy side* of favorite or underdog corresponding to the team chosen as more likely to beat the spread. This variable imposes the polarization precondition of having like-minded group members as the groups are homogeneous with respect to the fundamental policy question of which team will win the bet. The second variable is *disagreement level* of "high" or "low" that depends upon the difference between the minimum and maximum wagers in the group. Each group consisted of low, intermediate, and high wager individuals with respective wagers w_1, w_2, and w_3. In all groups, the intermediate wager was set so that $w_2 \in \{\$3, \$4\}$. In the high disagreement condition, $w_1 = \$0$ and $w_3 = \$7$ giving a difference of \$7. In the low disagreement condition $w_1 \in \{\$1, \$2\}$ and $w_3 \in \{\$5, \$6\}$ so that the difference could be \$3, \$4, or \$5. The third variable is *network structure* of "complete" in which all members could communicate with each other or "chain" in which the intermediate wager member w_2 served as the center node connecting w_1 and w_3. After discussion, each member made their final wager. A group decision was not required but groups arrived at a consensus wager far more often than the alternative outcomes of a two-person majority or three different wagers. A winning (losing) bet resulted in a payoff of \$7 plus (minus) the wager, which was donated to a charity.

Polarization, or more specifically a risky shift, is observed for a group if its mean wager after discussion is greater than its initial mean wager. Most of the 198 groups reached a consensus wager. For these 169 consensus groups, statistically significant results were observed for all three of the manipulated variables. For policy side, only the favorite side exhibited a risky shift whereas the underdog side did not. For disagreement level, restricted to favorite groups (as underdog groups showed no systematic risky shift), high disagreement groups exhibited a greater risky shift than low disagreement groups. For network structure, similarly restricted to favorites, complete networks showed a greater risky shift than chains. All three of these behaviors can be seen in Fig. 1 in which substantial polarization is observed when the error interval is above the initial mean.

The above results are not readily explained by standard polarization theory. Particularly challenging is the policy side result as standard theory predicts that both policy sides should show a risky shift. For persuasive arguments theory, members of both the favorite and underdog groups presumably possess novel information in support of their team choice and should therefore increase their confidence and wager. For social comparison theory, a norm toward risk taking should cause both sides to increase their wager. The extremist-tilting explanation prevalent in opinion network modeling also fails to explain this differential polarization by policy side: if individuals with more extreme wagers are taken to be more confident and persuasive,

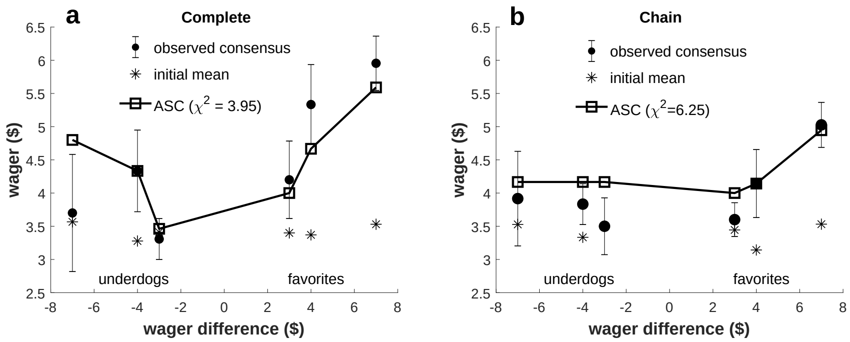

Fig. 1 Observed and simulated mean consensus wagers as function of initial wager difference, $w_3 - w_1$. (**a**) Complete network. (**b**) Chain network. Favorite groups shown on the positive x axis, underdogs on the negative. Observed consensus is average of final consensus wager (taken as positive for both favorites and underdogs) over groups at each difference value (no $5 difference groups were used as there were only four total). Also shown is average of the group mean initial wager. Experimental data shown as circles. Error bars are standard errors. χ^2 value is the sum of the squared errors between the simulated and the experimental values normalized by the standard error at each data point. Simulation parameters: $\alpha = 0.034$, $\lambda(0) = 0.03$, $\lambda_{min} = 0.01$.

then both favorite and underdog groups should display an equal tendency to increase their wagers.

4 Frame-Induced Polarization Theory

Reference [12] proposes a novel theoretical mechanism for group polarization that explains the results of the experiment. Central to the proposed mechanism is the distinction between the quantitative policy under debate and the *rhetorical frame* — the aspect of the policy upon which deliberations focus. The rhetorical frame will typically correspond to the dominant source of disagreement within the group due, for instance, to uncertainty as to the likelihood of an outcome. In a binary gamble such as in the experiment, the policy (e.g. wager amount) on a given outcome (e.g. team) and the rhetorical frame should be the subjective probability that that outcome will occur (e.g. win against the spread). The rhetorical frame position $\rho(x)$ is taken to be a function of the policy x. Groups will tend to shift toward the extreme if the functional relationship between the rhetorical position and the policy is concave ($\rho'' < 0$), that is, the rhetorical position increases more slowly as the policy becomes more extreme. For the experiment, such a concave relationship is expected between the subjective probability that a subject's chosen team will win against the spread and the wager amount (see Sec. 6).

The effect of concavity is to compress rhetorical distances toward the extreme relative to the distances between more moderate members, making it easier for majorities to form on the extreme side of the mean. Consequently, while the policy distribution may be symmetric so that no majority is favored on either side of the mean (as is approximately the case in our experiment), the distribution of rhetorical positions is skewed so that there is an initial majority on the extreme side of the rhetorical mean. This rhetorically-proximate majority (RPM) converges to a policy position more extreme than the mean to which the moderate minority of group members (those with policies below the mean) then concur, thereby resulting in a consensus policy that exhibits group polarization. The members of the F group (analogous to the favorite groups) in Fig. 2 provide an example of this mechanism. Although the intermediate member F_2 is equidistant in policy from the moderate F_1 and the extremist F_3, F_2 is rhetorically closer to F_3 and therefore (F_2, F_3) is the RPM pair. They agree on a policy halfway between them to which F_1 comes up due to majority influence. The RPM policy (b) is seen to be greater than the initial mean policy (a).

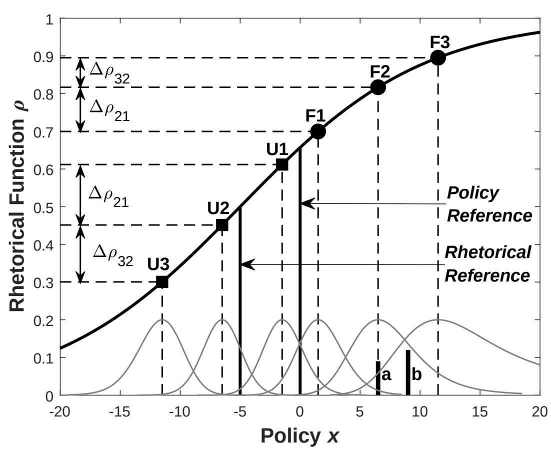

Fig. 2 Effect of rhetorical function concavity and offset reference.$\rho(x) = 1/(1 + e^{-\beta(x-x_0)})$ with $\beta = 0.13$, $x_0 = -5$. Short lines at bottom show alternative F group consensus policies: (a) mean policy $\bar{x} = x_2$; (b) RPM policy $\bar{x}_{23} = (x_2 + x_3)/2$. ASC model acceptance functions in gray.

Although the concavity of the rhetorical function explains the basic group polarization effect, it cannot by itself account for unequal polarization on opposing policy sides as observed in the experiment. Capturing this differential polarization involves the freedom of the rhetorical function to have a different reference point than the policy. The policy reference is defined as the neutral point, taken to be $x = 0$, that demarcates opposing policy sides. The rhetorical reference is defined as the policy value that maps to the neutral point of the rhetorical frame. For a *proper* frame, the rhetorical reference is the same as the policy reference so that the pro and con policy sides coincide with the pro and con rhetorical sides. For an *improper* frame, the rhetorical and policy references are offset so that the rhetorical reference splits one of the policy sides. Figure 2 shows how an improper frame can lead to differential polarization by policy side. The rhetorical reference splits the con (negative) policy side, which results in the U group (analogous to underdog groups) being arrayed on the approximately linear part of the rhetorical function rather than on the shoulder as for the F group. Consequently, U_2 is roughly the same rhetorical distance from both U_1 and U_3. Considering the effects of uncertainty and noise, formation of the moderate (U_1, U_2) RPM pair is about as likely as the extreme (U_2, U_3) pair so that systematic group polarization is absent or much reduced as observed in the experiment for the underdog groups. An improper frame can result from the heuristic substitution of a simpler, intuitive frame in place of a more complex proper frame that directly corresponds to the policy [12]. In the experiment, the heuristic frame of which team will win the *game* replaces the proper frame of who will win against the *spread*.

5 Accept-Shift-Constrict Model

The ASC model evolves both the positions and uncertainties of group members in response to their dyadic interactions. We consider position first, which can be a policy or, more generally, an opinion about some matter. A persuasive message sent by one group member to another must first be accepted by the recipient in order to shift their policy. While a number of factors can affect whether a message is accepted, the distance between the message's rhetorical position and that of the receiver plays the key role in our model: if the distance is within the *latitude of acceptance* (LOA), the message is likely to be accepted, but the acceptance probability rapidly decays beyond the LOA. If the message is accepted, then the receiver's policy is shifted in proportion to its distance from the sender's policy.

Formally, we encode the above process as an ordinary differential equation for $x_i(t)$, the policy position of the i^{th} group member at time t. For a group with N members, the rate of change of x_i is given by

$$\frac{dx_i}{dt} = \sum_{j=1}^{N} \nu_{ij}(x_j - x_i) \exp\left\{-\frac{1}{2}\frac{(\rho(x_j) - \rho(x_i))^2}{\lambda_i^2}\right\}, \tag{1}$$

where ν_{ij} is the coupling strength from $j \to i$ and λ_i is i's LOA. The matrix formed by the coupling strengths defines a position-independent network of influence. In general, ν_{ij} depends on communication rate and other factors such as credibility and expertise ($\nu_{ii} = 0$).

The linear $x_j - x_i$ term in Eq. (1) represents the shift effect. The gaussian term represents the acceptance process and we refer to it as the acceptance function, $a(\Delta\rho, \lambda) = e^{-\Delta\rho^2/2\lambda^2}$. Although the acceptance function is always symmetric with respect to the sign of the rhetorical difference, $a(-\Delta\rho) = a(\Delta\rho)$, a concave $\rho(x)$ can causes it to appear asymmetric along the policy axis as clearly seen for F_2 and F_3 in Fig. 2.

In addition to position change, communication can also affect a person's uncertainty regarding their position. Group discussion has been observed to increase the level of certainty that members have in their quantitative judgments [15]. Accordingly, we introduce an uncertainty reduction mechanism in our model in which messages from those with similar positions constrict an individual's LOA so that they become more resistant to persuasion from distant positions. Messages originating within the LOA that are accepted decrease the LOA, but not beneath a certain minimum value λ_{min}. This yields for the LOA dynamics:

$$\frac{d\lambda_i}{dt} = \begin{cases} \sum_{j=1}^{N} \nu_{ij}(\lambda_{min} - \lambda_i)e^{-\Delta\rho_{ij}^2/2\lambda_i^2}, & |\Delta\rho_{ij}| \leq \lambda_i \\ 0, & |\Delta\rho_{ij}| > \lambda_i. \end{cases} \tag{2}$$

Equations (1) and (2) comprise the ASC model. Assuming no difference between rhetorical and policy positions, i.e, $\rho(x) = x$, Eq. (1) is equivalent to the model of [8] without the self-influence force that models a persistent effect of an individual's initial opinion. The uncertainty reduction dynamics represented by Eq. (2) is novel in opinion network modeling. The model of [13] includes a dyadic uncertainty interaction that results in uncertainty change only when dyad members have different uncertainties; this requires that uncertainty levels be visible to other group members, an assumption not present in Eq. (2), and does not allow equally uncertain individuals to mutually reinforce their opinions.

A crucial consequence of the uncertainty reduction dynamics in the ASC model is the ability for interim majorities to more effectively maintain their position in the face of minority influence. This effect is essential to the RPM process in the theoretical account of group polarization above but it occurs regardless of whether or not the rhetorical function is different from the policy. Figure 3a illustrates the rough persistence of the majority position for a complete-network triad in which the intermediate member's position is taken to be halfway between the others. For sufficiently low initial disagreement, however, an interim majority will not form and the group equilibrium will be close to its initial mean (Fig. 3b).[2]

6 Simulation of Group Polarization

This section demonstrates the ability of the ASC model to produce the same qualitative effects as in the frame-induced polarization theory and as observed experimentally. Going beyond qualitative correspondence, its agreement with the data on a quantitative level is also shown. First, we discuss how the coupling strengths ν_{ij} are set. They are treated as dyadic communication rates as determined by simple

[2] The persistence of majority positions on a continuous opinion axis is also found in the agent-based model of [16], which employs a confidence variable that must be transmitted between agents along with opinions, rather than the ASC model's use of an uncertainty interval not visible to others.

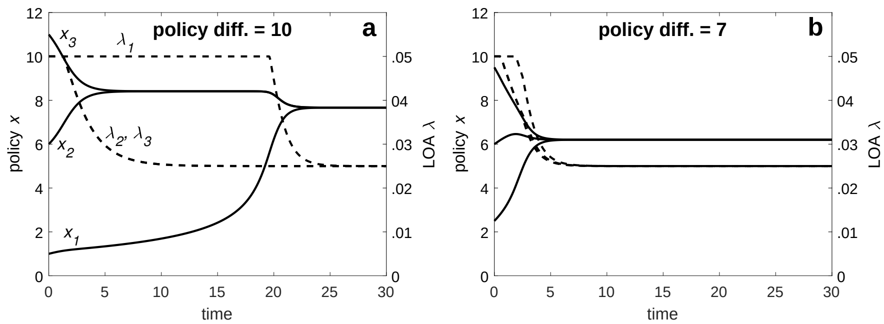

Fig. 3 Position and LOA trajectories in ASC model for a complete network. Solid curves show policy positions, dashed curves show LOAs. (**a**) High initial policy disagreement ($x_3 - x_1 = 10$) showing substantial shift between consensus and initial mean policy ($x_2(0)$). (**b**) Lower initial policy disagreement (7) results in near simultaneous convergence close to initial policy mean. $\lambda_{1,2,3}(0) = 0.05$, $\lambda_{min} = 0.025$; $\rho(x)$ as in Fig. 2.

topological considerations. For a complete network, on average, the communication rates are expected to be the same for all nodes, so we set $\nu_{ij} = 1/2$ for all three dyads. For the chain, if the sequence in which nodes send messages follows the chain path and the center node (node 2) predominantly opts to send its messages simultaneously to both outer nodes (rather than separately), then we expect node 2 to have about twice the communication rate with each of nodes 1 and 3. We therefore set $\nu_{12} = \nu_{32} = 1$ and $\nu_{21} = \nu_{23} = 1/2$.[3] These communication rate expectations are approximately borne out in our experiment.

Figure 4 displays simulation results for complete and chain network triads that are homogeneous with respect to policy side analogous to the experimental setup. The baseline case (dotted curve) consists of an intermediate node with an initial policy $x_2(0)$ halfway between the initial positions of the moderate $x_1(0)$ and the extremist $x_3(0)$. The other cases shown (light gray curves) account for position uncertainty by allowing $x_2(0)$ to deviate by various small amounts from the baseline case. The discussion-induced shift in the mean is plotted against the initial policy difference between the extremist and the moderate, where the opposing pro and con policy sides are shown on the positive and negative sides of the horizontal axis respectively. For the pro (con) side, a positive (negative) polarization shift indicates a shift toward

[3] The sum of the communication weights is normalized to the same (arbitrary) value of 3 in both networks, a value that only affects the transient time and not the final equilibrium.

the extreme — a higher wager in the case of the experiment. The mean over all the cases (solid dark curve) can be used to gauge the extent of systematic polarization.

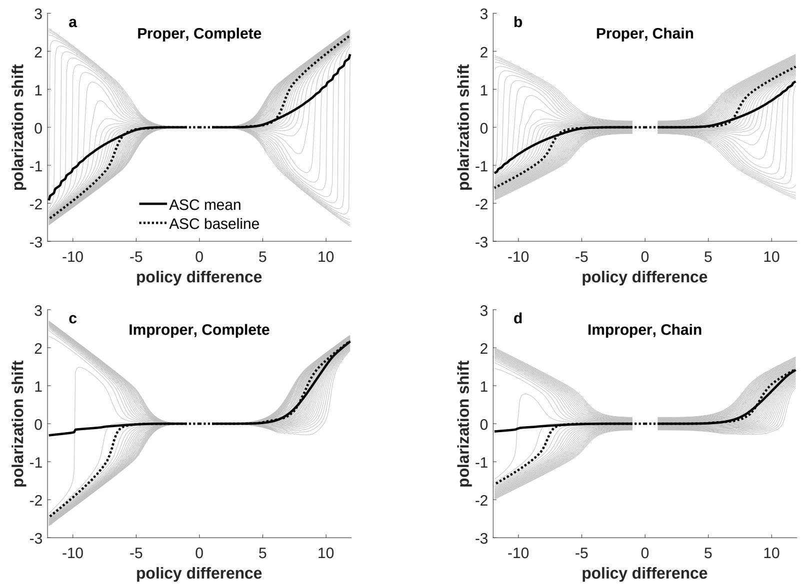

Fig. 4 ASC simulations for triad networks with variability in intermediate node policy. $\rho(x)$ taken as in Fig. 2. Top row shows proper rhetorical frame ($x_0 = 0$). Bottom row shows improper rhetorical frame ($x_0 = -5$). Positive and negative policy sides are on positive and negative horizontal axis respectively. Polarization shift, $\delta = \bar{x}(t_f) - \bar{x}(0)$, plotted as a function of the initial policy difference, $\Delta = x_3(0) - x_1(0)$. Shift toward the extreme corresponds to $\delta > 0$ for positive policy side and $\delta < 0$ for negative side. The position of the intermediate node was varied according to $x_2(0) = \pm(6 + \epsilon)$ for the positive and negative policy sides, where ϵ takes on 41 uniformly-spaced values over the interval $[-1, 1]$. $x_1(0) = 6 - \Delta/2$ and $x_3(0) = 6 + \Delta/2$ for $\Delta > 0$ and analogously for $\Delta < 0$. ASC mean (black) taken over all ϵ values. Shifts for individual ϵ values shown as gray curves. Dotted curve shows $\epsilon = 0$ baseline case where $x_2(0) = 6$. Gap in the curves is the region where $x_2(0)$ would go beyond $x_1(0)$ or $x_3(0)$. ASC model parameters: $\lambda_{1,2,3}(0) = 0.05$, $\lambda_{min} = 0.025$.

The top row of Fig. 4 represents a proper rhetorical frame in which the policy and rhetorical references are coincident. In the experiment, the proper frame is the subjective probability of the favorite winning against the spread. The rhetorical function is taken to be concave with increasing policy extremity.[4] Regarding the mean, both policy sides exhibit equal polarization that increases with disagreement level and with the complete network showing more polarization than the chain. Considering higher disagreement levels, the mean polarizes less than the baseline case because some groups actually depolarize — those in which the moderate and

[4] If the subjective probability of one of the binary outcomes is taken as the rhetorical frame and opposing policy sides have opposite signs, then concavity with increasing policy extremity yields an overall S-shaped rhetorical function as explained in [12].

intermediate node are sufficiently close to overcome the skewing effect of the rhetorical function. This ability to predict depolarization for individual groups despite the dominant tendency toward polarization is an important capability not present in the informational, normative, or extremist-tilting theories. Although the proper frame does exhibit polarization, accounting for the differential polarization by policy side observed experimentally requires use of an improper frame as is the subjective probability that the favorite will win the game. The bottom row of Fig. 4 employs an improper frame and indeed shows substantial polarization for positive policies and little for negative ones.

The ASC model can also be quantitatively tested against the data. Groups can be simulated using their actual initial wagers and with the communication weights as set above. The rhetorical function $\rho(w)$ that maps the wager to the subjective probability of a favorite game victory (the improper frame) is derived in [12] based on the theory of individual decision making under risk and uncertainty. It depends upon the subjective probability $p(w)$ of a favorite victory (the proper frame)

$$p(w) = \frac{1}{2} - \frac{1}{8\alpha w} \pm \frac{1}{2}\sqrt{1 + \frac{1}{16\alpha^2 w^2}}, \tag{3}$$

where the + (−) sign implies bets on the favorite (underdog). The free parameter α is the risk aversion that quantifies how sensitive individuals are to variance around the expected value of the payoff. It is assumed to be identical for all subjects. The rhetorical function is then given by

$$\rho(w) = \frac{1}{2}\mathrm{erfc}\left\{\mathrm{erfc}^{-1}\left(2p(w)\right) - \frac{s_0}{\sigma\sqrt{2}}\right\}, \tag{4}$$

where $\mathrm{erfc}(u) = \frac{2}{\sqrt{\pi}}\int_u^\infty e^{-v^2}dv$. The parameter s_0 is the point spread for the game in question and $\sigma = 12.8$ is the empirical standard deviation for the margin of victory in NFL games. Both $p(w)$ and $\rho(w)$ are S-shaped implying a concave relationship between the subjective probability of the outcome estimated as more likely and the wager magnitude.

In addition to the risk aversion, there are two free parameters from the ASC model that need to be fit from the data, the initial LOA $\lambda(0)$ and the minimum LOA λ_{min}, both assumed identical for all subjects. The three parameters are estimated by minimizing the sum of χ^2 error values over both complete and chain networks. The simulation results are shown in Fig. 1. A three-parameter χ^2 goodness-of-fit test, which takes as its null hypothesis that the model is correct, yields a probability $Q = 0.33$ that χ^2 could have exceeded its observed value of 10.2 by chance. With a conservative threshold of $Q < 0.2$ for rejecting the null hypothesis, the ASC model is found to be consistent with the data.

7 Conclusion

The ASC model presented here describes a dual process of opinion and uncertainty change based on the greater acceptance rate of messages within one's LOA and the decrease in LOA due to exposure to similar views. A key dynamic in the model is the ability of proximate majorities to form and persist for symmetric networks, thereby enabling majorities to exert outsized influence and produce a consensus opinion different from the initial mean. Importantly, the ASC model does not involve the exchange of uncertainties over the network unlike other models in which uncertainties are directly coupled along with opinions [13, 16]. Another important innovation of the ASC model is the conceptualization of distinct dimensions of opinion and rhetoric: opinion is an evaluation directly tied to a decision or other behavioral outcome of interest while rhetoric determines whether messages aimed at shifting opinions are found persuasive. If the rhetorical function mapping opinion to rhetorical position is concave, then proximate majority formation at the extreme is facilitated. Consequently, the ASC model can generate systematic group polarization due to the structure of the decision space rather than by assuming an asymmetric network structure in which influence is associated with extremity as typically done in opinion network modeling. The ASC model simulations shown here display the same qualitative phenomena as observed in the experiment: polarization on one policy side but not the other, increasing polarization with disagreement level, and greater polarization for complete networks than for chains. Furthermore, the ASC model is in quantitative agreement with the experimental data.

References

1. C. Castellano, S. Fortunato, V. Loreto, Reviews of Modern Physics **81**(2), 591 (2009)
2. D. Kempe, J. Kleinberg, S. Oren, A. Slivkins, Network Science **4**(01), 1 (2016)
3. A.V. Proskurnikov, R. Tempo, Annual Reviews in Control **43**(Supplement C), 65
4. M.H. DeGroot, Journal of the American Statistical Association **69**(345), 118 (1974)
5. N.E. Friedkin, E.C. Johnsen, *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics* (Cambridge University Press, Cambridge, UK, 2011)
6. R. Olfati-Saber, J.A. Fax, R.M. Murray, Proc. IEEE **95**(1), 215 (2007)
7. J. Lorenz, International Journal of Modern Physics C **18**(12), 1819 (2007)
8. M. Gabbay, Physica A **378**, 118 (2007)
9. D.G. Myers, H. Lamm, Psychological Bulletin **83**(4), 602 (1976)
10. D.J. Isenberg, Journal of Personality and Social Psychology **50**(6), 1141 (1986)
11. C.R. Sunstein, Journal of Political Philosophy **10**(2), 175 (2002)
12. M. Gabbay, Z. Kelly, J. Reedy, J. Gastil, Social Psychology Quarterly **81**(3), 248
13. G. Deffuant, F. Amblard, G. Weisbuch, T. Faure, Journal of Artificial Societies and Social Simulation **5**(4) (2002)
14. N.E. Friedkin, Control Systems, IEEE **35**(3), 40 (2015)
15. J.A. Sniezek, Organizational Behavior and Human Decision Processes **52**(1), 124 (1992)
16. M. Moussaid, J.E. Kammer, P.P. Analytis, H. Neth, PLOS ONE **8** (2013)

Acknowledgements This work was supported by the Office of Naval Research under grant N00014–15–1–2549.

Appendix 2

Networks of Cooperation: Rebel Alliances in Fragmented Civil Wars

Networks of Cooperation: Rebel Alliances in Fragmented Civil Wars

Emily Kalah Gade,
Department of Political Science, University of Washington

Michael Gabbay,
Applied Physics Laboratory, University of Washington

Mohammed M Hafez,
Department of National Security Affairs, Naval Postgraduate School

Zane Kelly,
Applied Physics Laboratory, University of Washington

Abstract

When rebels make alliances, what informs their choice of allies? Civil wars are rarely simple contests between rebels and incumbent regimes. Rather, rival militant networks provide the context in which these fragmented conflicts unfold. Alliances that emerge within this competitive landscape have the power to alter conflict trajectories and shape their outcomes. Yet, patterns of interrebel cooperation are understudied. The existing scholarship on rebel alliances focuses on why rebels cooperate, but little attention is given to the composition of those alliances: with whom rebels cooperate. We explore how power, ideology, and state sponsorship can shape alliance choices in multiparty civil wars. Employing network analysis and an original dataset of tactical cooperation among Syrian rebels, we find compelling evidence that ideological homophily is a primary driver of rebel collaboration. Our findings contribute to an emerging literature that reasserts the role of ideology in conflict processes.

Keywords: Syria, civil war, fragmentation, alliances, social network analysis, ideology

Corresponding author contact: ekgade@uw.edu

Civil wars are rarely simple contests between unified rebels and incumbent regimes. Instead, they usually feature divided rebel movements with multiple factions competing over leadership, territory, resources, and fighters (Bakke, Cunningham, and Seymour 2012; Cunningham 2011). Forging unity among armed groups is a challenge because credible commitment problems make binding obligations difficult to initiate and sustain. Cooperation also involves tradeoffs between enhancing one's power capabilities and decision-making autonomy, which may incline some rebels to forgo alliances that diminish their independence. Ideological considerations also affect rebel cooperation: factions that harbor competing visions for the future are likely to view alliances with rivals as short-lived exigencies at best. Yet despite these barriers, cooperation among armed factions does occur. Between 1946 and 2008, 181 out of 345 groups in civil wars, more than 52 percent, "have initiated positive associations with each other while fighting with the government" (Akcinaroglu 2012, 890).

The prevalence of rebel cooperation alongside competition generates two puzzles: why do rebels cooperate, and with whom do they cooperate? Literature on rebel cooperation has focused on the *why* question. Interrebel alliances emerge between factions seeking to augment their capabilities and improve their tactical productivity (Lichbach 1995), balance against their rivals through minimum winning coalitions (Christia 2012), and increase their overall odds of victory by institutionalizing joint command and control of military operations (Akcinaroglu 2012). Little is known, however, about the factors that shape the composition of rebel alliances, i.e. *with whom* rebels cooperate. Civil wars can involve hundreds of rebel brigades, which could produce countless cooperative alignments. This translates into rebels having choices when pursuing cooperation to achieve their conflict objectives. What explains their choice of allies?

In addressing this puzzle, we make distinct theoretical, methodological, and empirical contributions. Theoretically, we explore three logics of alliance composition related to ideology, power, and state sponsorship and make predictions about how they might shape militant collaboration. We posit that ideological proximity in rebel networks should yield greater militant cooperation than ideological distance, thus challenging the prevailing assumption that ideology is a minor consideration in alliance formation (Christia 2012). We operationalize ideology in civil wars along three dimensions—*conflict framing*, *conception of the ideal polity,* and *territorial aspiration*—and show that agreement within those issue areas facilitates cooperation among rebel factions. Through conflict framing, a rebel group identifies whom it is primarily fighting for and against, casting both ingroup and outgroups with respect to its preferred cleavage whether ethnic, religious, economic, or political. A group's conception of the ideal polity identifies its vision for the post-conflict social and political order, and its territorial aspiration identifies the boundaries of this future order. Unpacking ideology into these distinct dimensions allows for a more nuanced understanding of factional alignments than the classification of rebels into broad categories such nationalists, separatists, socialists, and fundamentalists.

We complement our ideological analysis with a thorough consideration of how power and state sponsorship inform alliance choices. For power, we propose contrasting hypotheses of symmetrical and asymmetrical alliance formation. We posit that an overriding concern for capability aggregation in rebel movements will tend to produce *symmetric alliances* (cooperation between groups of comparable strength), whereas the desire of strong groups to form alliances that maximize decision-making autonomy vis-à-vis rivals will generate *asymmetric alliances* (cooperation between groups of dissimilar power capabilities). As for state sponsorship in rebel alliances, we test the hypothesis that rebel groups that share the same state sponsor will cooperate more frequently than those with no overlapping external sponsors.

Methodologically, we employ social network analysis to yield insights into rebel cooperation within fragmented conflicts. Network approaches are widespread in political science, yet few have sought to apply them to multiparty civil wars (Zech and Gabbay 2016; Metternich et al. 2013). Research on civil conflicts calls for a network approach because rebel groups do not make choices to align with others in a vacuum, but rather their choices are likely to hinge on the alliance preferences of the other groups in the rebel movement. Thus, social network analysis can better capture the theoretical patterns we would expect to observe than the standard statistical assumption of independence of observations when examining dyads in multiparty civil wars. We use Additive and Multiplicative Effects (AME) models to evaluate the relationships between our three proposed variables simultaneously. In addition, as a robustness check, we use simulations of network tie formation to augment these findings.

Empirically, we test our hypotheses as they relate to factional cooperation in Syria's civil war. One of the world's bloodiest conflicts, the Syrian rebellion features a complex set of actors with local, regional, and international ties. We use primary insurgent sources, including more than 9000 unique claims of attacks, to construct an original dataset for more than 220 insurgent groups active since the onset of conflict through mid-2015. We form a network of militant tactical cooperation from claims of joint operations and investigate its structure with respect to ideology as obtained from manual coding of primary source materials, power as measured by group size, and state sponsorship as reported in informed secondary sources. We find compelling evidence that ideological homophily is a driver of rebel cooperation in Syrian militant networks. We also find some evidence in favor of symmetric alliances rather than asymmetric ones, but it is inconsistent across our analysis. However, we do not find support for the proposition that overlapping state sponsorship in rebel dyads increases cooperation.

The question of alliance composition has important strategic implications. Understanding the dynamics of rebel cooperation can yield policy insights for prompting or dissuading alliances between rebel groups in multifactional civil wars. Recent conflicts illustrate vividly how the composition of rebel alliances can shape conflict trajectories in dramatic ways. In 2007, American-led coalition forces in Iraq successfully exploited rifts in rebel unity to turn the tide in the war. Nationalist insurgents became increasingly alienated by Al-Qaeda in Iraq, their former jihadist ally, and as such were willing to switch sides to the American coalition. The dissolution of the nationalist-jihadist alliance contributed to a substantial reduction in violence, until the resurgence of the Islamic State six years later.

In contrast, the United States viewed with concern the fragmentation of Syria's rebel movement, preferring a unified and cohesive rebel alliance composed of moderate rebel factions that could topple the Assad regime. The United States, however, could not overlook the presence of extreme Islamist factions in the rebel movement, and thus deprioritized the objective of rebel unity. Ultimately, it limited its support to a narrow sector of acceptable militant groups, which proved ineffective against an incumbent regime backed by a unified and powerful coalition consisting of Iran, Hezbollah, and Russia.

Beyond these recent examples, the literature on rebel fragmentation points out that (dis)unity has important implications for several other conflict processes. Civil wars with divided rebel factions last longer, are more violent, and have higher rates of recurrence than wars with unified rebel movements (Cunningham, Bakke, and Seymour 2012; Wood and Kathman 2015; Driscoll 2012; Cunningham 2013; Rudloff and Findley 2016). Conversely, movements led by a hegemonic faction are more likely to be successful than more diffuse movements (Krause 2017). By illuminating the drivers of rebel cooperation, this study, therefore, makes a contribution to understanding a dynamic of great consequence in fragmented conflicts.

Ideology in Rebel Alliances

Rebel groups have political preferences and moral visions for which they are fighting. The preceding century has highlighted the capacity of Marxist, nationalist, fundamentalist, and fascist ideologies to mobilize millions of people for revolutions, insurgencies, civil wars, and genocide. Although not all civil conflicts are driven by ideological divides and not all rebels are motivated by ideological considerations, diversity of political demands typify fragmented civil conflicts, which are the most common form of wars today (Seymour, Bakke, and Cunningham 2016, 5-6; Jones 2017, 168).[1]

Recent scholarship has rediscovered the critical role that ideology plays in conflict processes (Ugarriza and Craig 2012; Costalli and Ruggeri 2015; Staniland 2015; Balcells 2017). Ideology is a source of collective identity and can help forge group cohesion in the context of civil wars by orienting commanders and foot soldiers toward a clear set of objectives (Gutiérrez Sanín and Wood 2014). It can also motivate commitment and sacrifice, remove inhibitions to violence, and reprioritize collective incentives above self-regarding considerations (Lichbach 1995, 92-93; Walter 2017, 19-20; Kim 2018, 308). Additionally, ideological socialization has been shown to improve battlefield discipline and dissuade defections to the state (Oppenheim, et al. 2015; Hoover Green, 2016).

We contribute to this burgeoning literature by proposing mechanisms that link ideology to the choice of allies in rebel movements, thus challenging the prevailing assumption in the literature that ideology is a secondary consideration in alliance formation. We expect these mechanisms to apply to the spectrum of cooperative relationships ranging from joint operations at the tactical level to formal alliances at the strategic level. Joint operations (our empirical measure) consist of two or more rebel groups conducting an attack together (Bapat and Bond 2012, 19). We focus here on tactical collaboration because only 17.6 percent of rebel cooperation between 1946 and 2008 was at the level of formal alliances (Akcinaroglu 2012, 890).

Joining Gutiérrez Sanín and Wood (2014, 215), we define ideology as "a more or less systematic set of ideas that includes the identification of a referent group (a class, ethnic, or other social group), an enunciation of the grievances or challenges that the group confronts, the identification of objectives on behalf of that group (political change – or defense against its threat), and a…program of action." We operationalize this definition by disaggregating ideology along three axes: *conflict framing, conception of the ideal polity,* and *territorial aspiration*. Each of these dimensions suggests causal mechanisms that link ideology to rebel alliances.

Conflict framing refers to how rebel factions demarcate the core political, religious, or social categories that constitute one's ingroup and outgroups.[2] A group's conflict frame specifies its preferred conflict dyad, the outgroup most threatening to the ingroup. In Iraq, for example, nationalist insurgents opposed to America's 2003 invasion of their country employed a resistance frame of Iraqis versus American occupation forces as their primary conflict frame; the Iraqi government and Shiite militias were viewed as mere instruments of America's occupation. In contrast, jihadist groups, especially Al-Qaeda in Iraq, framed the conflict as a sectarian struggle between Sunnis and Shiites, whereby American forces enabled Shiites to dominate Sunnis. In each instance, conflict framing implies that threats from a particular outgroup are more salient than others, and that certain parties to the conflict could conceivably cooperate while others are unthinkable; Sunni nationalists could ally with Shiites whereas sectarian jihadist groups could not. Thus, the conflict frame ingroup bounds the choice of allies.

As a group's conflict frame helps determine whom it attacks, conflict framing may also indirectly promote cooperation to the extent that groups with similar targeting portfolios can more easily cooperate. For example, two rebel groups – one nationalist and one sectarian – may both primarily target the state's security forces, yet the first casts them as the goons of a tyrannical regime while the second casts those same forces as the soldiers of the rival sect. Although the pair could cooperate on the basis of this common targeting, if the sectarian framing is also extended to justify indiscriminate and controversial attacks against rival sect civilians (included within the nationalist ingroup), then the associated dissension would inhibit cooperation.

Conception of the ideal polity is the normatively prescriptive dimension of ideology that orients members to a vision of the desired end state. It specifies how groups define a legitimate sociopolitical order that is worth fighting for, deeming some institutional arrangements appropriate while viewing others as unjust, inequitable, oppressive, or even heretical. This dimension captures the traditional ideological divides such as the competition between the economic left and right, democrats and authoritarians, and secularists and fundamentalists. When choosing to form alliances, we expect rebels to align with those that offer mutual political corroboration and are working toward similar objectives. Groups with fundamentally divergent post-conflict goals or territorial aspirations will have a greater ideological distance to traverse in order to achieve cooperation.

Territorial aspiration delineates the boundaries of the ideal state and orients rebels to the territorial claims of the movement. This dimension captures the degree to which rebels seek to maintain or violate the territorial integrity of their states.[3] Movements with shared conceptions of the ideal polity sometimes diverge over the territorial boundaries of that polity. For example, parties representing Basques and Catalans in Spain diverge on the issue of maintaining local autonomy or insisting on separatism as do Scots in the United Kingdom. Arab nationalists in their heyday were divided between advocates of *wataniyya* (homeland patriotism) and *qawmiyya* (pan-Arab unification). Islamists today are divided between those who favor establishing an Islamic order within the framework of the modern national state, and those that harbor the irredentist ambition of restoring an Islamic caliphate.

Like the previous dimensions, territorial aspiration is a potential source of unity or division. Separatist groups, for example, may be unwilling to compromise their own territorial demands, creating friction with nationally-focused groups. Territorial aspirations are likely to accentuate disagreements between factions as a conflict becomes protracted. Groups that care about the territorial integrity of their states may incline toward a negotiated end to the conflict in order to restore national unity. Those that harbor broader territorial ambitions are less likely to prioritize national unity as the conflict persists and may be inclined to sabotage conflict-ending negotiations.

Agreement on conflict framing, ideal polity, and territorial aspiration, therefore, predict *ideological homophily* in network ties. A fundamental principle of social network analysis, homophily states that "similarity breeds connection," and social networks tend to be largely homogenous because ties between dissimilar individuals dissolve more quickly (McPherson, Smith-Lovin, and Cook 2001, 415-416). Homophily prevails because of the presumption of mutual trust and complementarity of interests among actors with uniform attributes (Lichbach 1995, 138-141), and because joining similar others reinforces the cognitive bias toward belief confirmation in polarized political contexts (Balliet et al. 2016, 15-16). Political homophily has been observed at the individual, organizational, and state levels, including life style politics (DellaPosta et al. 2015), online activism (Boutyline and Willer 2017), local government regional planning networks (Gerber, Henry, and Lubell 2013), international trade networks (Maoz 2012), third party state interventions (Corbetta 2013), and international alliances (Werner and Lemke 1997; Lai and Reiter 2000; Wendel 2001). We anticipate ideological homophily will also shape rebel alliance choices, yielding our ideology hypothesis:

H_1: Ideological alliances: Interrebel cooperation is more likely among ideologically similar groups than ideologically dissimilar ones.

Complementary Logics of Alliance Formation

We consider two logics of alliance composition based on power and state sponsorship, which may operate in conjunction with ideological homophily.

Power in Rebel Alliances

In the most extensive analysis of rebel cooperation, Christia (2012, 240) advances the power-centric theory of minimum winning coalitions (MWC), which are "alliances with enough aggregate power to win the conflict, but with as few partners as possible so that the group can maximize its share of postwar political control." Absent credible commitments, however, weaker alliance members grow wary of their stronger partner as the alliance nears victory. Hence, the theory predicts coalitional instability as rebels regularly switch sides, thereby maintaining a rough balance of power. Apart from this balancing constraint, the theory remains silent on the composition of the rival coalitions. To the extent that the MWC theory considers the credible commitments problem at its utmost severity, it expects little association between ideology and militant cooperative relationships (Christia 2012, 32-33).

We propose two contrasting hypotheses about how relative power considerations may affect alliance composition beyond balancing constraints. The first relates to *symmetric alliances* (cooperation between groups of similar power capabilities). Rebel groups in search of greater security may form alliances to aggregate their capabilities against mutual threats.[4] Given that the pooling of assets and coordination of tactics becomes more difficult as the number of groups grows, two powerful groups can cooperate more efficiently than a coalition consisting of a powerful group and multiple minor groups.[5] If powerful factions prefer to coordinate with each other, that leaves weak factions to ally with other minor players. Thus, our first power hypothesis predicts:

H_2: Symmetric alliances: Interrebel cooperation motivated by mutual security concerns will produce cooperative network ties between groups of comparable power capabilities.

Rebel groups may also seek to maximize their decision-making autonomy in addition to augmenting their capabilities through *asymmetric alliances*—cooperation between major and minor rebel groups.[6] Groups that do not feel particularly threatened by the regime may prioritize winning on their own terms. Powerful rebels, in particular, can afford to emphasize enhanced autonomy over security by forming alliances with weaker partners amenable to influence. The weaker faction receives greater security from its alignment with a powerful group, while the dominant rebel faction benefits from both capability aggregation and control over the conduct of minor groups. Thus, we hypothesize:

H_3: Asymmetric alliances: Interrebel cooperation motivated by security and autonomy considerations will produce cooperative network ties between groups with dissimilar power capabilities.

State Sponsorship in Rebel Alliances

External sponsorship of rebel movements is a common feature of civil wars.[7] Rebels covet military, financial, and political support to outmatch the resources of their incumbent regimes, establish international legitimacy, exercise leverage in negotiations, and outcompete rivals. As Gurr (1970, 269) observed long ago, "The greatest potential increment to dissident military capacity is external support." Indeed, Jones (2017: 136) finds that insurgent movements that receive great power support win nearly half to two-thirds of the time.

External patrons provide arms, money, supplies, or sanctuaries to rebel groups in the expectation that these rebels will exhibit sufficient discipline and solidarity to fulfil their patron's strategic aims (Salehyan, 2010). Bapat and Bond (2012) and Popovic (2018) view external leverage as an important interrebel institution that can help overcome the credible commitments problem, police against side negotiations, and mediate conflicts between rebel groups. This predicts greater interrebel cooperation because sponsors can threaten to withhold financing and war materiel from those who are jeopardizing a cohesive rebel coalition (Lichbach, 1995: 179).

However, state sponsors can also undermine rebel unity by incentivizing some rebels to challenge their rivals (Tamm, 2016). This is particularly the case when multiple state sponsors with competing political agendas seek to foster their own proxy clients through patronage. The presence of multiple sponsors increases the number of avenues rebel groups have to support themselves and reduces the leverage any individual external patron can exert to foster cohesive rebel coalitions (Salehyan, Siroky, and Wood, 2014).

Acknowledging these contradictory effects of state sponsorship on rebel alliances, we propose that two rebel groups that share a single sponsor are more likely to cooperate with one another than dyads with distinct sponsors. This yields our final hypothesis:

H_4: State sponsored alliances: Rebel groups that derive support from the same state sponsor will experience greater cooperation than those lacking a common state sponsor.

Table 1 summarizes our hypotheses and suggests their observable implications for the composition of alliance networks at the level of tactical joint operations.

[Table 1 about here]

Table 1. Hypotheses, causal mechanisms, and expected network structure

Hypotheses	Causal Mechanisms	Expected Network Outcomes
H_1: Ideological Alliances	*Ideological homophily* shapes cooperation due to similar understanding of enemies and allies (conflict framing), ideas and institutions of sociopolitical order (ideal polity), and the boundaries of that order (territorial aspiration).	Network structure will be shaped by groups' shared ideological attributes.
H_2: Symmetrical Alliances	*Power aggregation* is the primary consideration behind cooperation.	Network structure will be shaped by groups' comparable power.
H_3: Asymmetrical Alliances	*Security-autonomy tradeoff* is the primary consideration behind cooperation.	Network structure will be shaped by groups' disparate power.
H_4: State Sponsored Alliances	*Sharing state sponsors* compels rebels to forge cohesive alliances.	Network structure will be shaped by groups' shared state sponsors.

Network Analysis of Syrian Militant Alliances

We employ social network analysis to test our four theoretical propositions. A social network consists of nodes and the ties between node dyads. The nodes can represent individuals, organizations, or states and ties can correspond to relationships such as communication, cooperation, and conflict. Social network analysiscan account for the interdependence of relationships within a set of political actors (Ward, Stovel, and Sacks 2011; Hafner-Burton, Kahler, and Montgomery 2009). Alliance models that assume independence of observations in dyads miss out on relational data because alliances are not created in a vacuum; they are dependent upon relationships with multiple groups (Dorussen, Gartzke, and Westerwinter 2016).

The fragmented nature of civil conflicts implies network analysis should be a fruitful methodology for addressing militant behaviors, such as alliance formation and infighting (Zech and Gabbay 2016). We illustrate the utility of the network approach in our analysis of rebel alliances in the Syrian civil war.

Rebel Factions in the Syrian Civil War

In March 2011, Arab Spring protest waves reached Syria after making their way from North Africa to the Middle East. Bashar al-Assad's regime initially sought to quell protests and prevent their diffusion through a mix of repression and concessions. However, these measures failed as protests gained momentum across Syria's major cities, and the protestors' demands shifted from reforms to regime change. As the conflict became militarized, the Free Syrian Army (FSA) formed from the ranks of defecting officers and its affiliated brigades began engaging in conventional armed attacks against regime forces. The FSA exemplified the secular nationalist tendency, framing the Syrian rebellion as a national and democratic revolution that encompasses Syria's diverse ethnic and religious communities.

The inability of protesters and the FSA to topple the Assad regime in the opening months of the insurgency gave rise to rival armed factions, the most notable of which was the Al-Qaeda affiliated Al-Nusrah Front (ANF). Formed in January 2012, ANF was avowedly sectarian and jihadist, casting the conflict not as a revolution, but rather as a holy war against a secular regime dominated by heretical Alawites (an offshoot sect of Shiism). It called for the formation of an Islamic state strictly adherent to religious law (International Crisis Group 2012, Lister 2015).

Many other Islamist factions emerged, ranging from Muslim Brotherhood sympathizers such as Al-Tawhid Brigade (ATB) to Salafists such as Ahrar al-Sham Islamic Movement (ASIM). The latter became one of the dominant factions in the insurgency, competing with both the FSA and ANF (International Crisis Group 2012). ASIM represented Salafist nationalists that wanted to establish an Islamic state within the boundaries of Syria's national territory, but, unlike ANF, it did not frame the conflict in sectarian terms.

Kurdish communities established their own armed groups, notably the People's Protection Units (Yekîneyên Parastina Gel, YPG), to defend their territories from regime forces as well as hostile rebels (International Crisis Group 2014). The YPG is secular in orientation and views Kurdish co-ethnics as its primary ingroup for whom it seeks autonomy within, or secession from, the Syrian state.

In 2013, the Islamic State in Iraq and the Levant (ISIL), led by Abu Bakr al-Baghdadi, splintered the ranks of ANF to form an even more extreme sectarian jihadist faction. ISIL further aggravated the conflict by intensifying sectarian polarization, expanding the conflict into neighboring states, and threatening international security through global terrorism. ISIL also produced fratricidal violence within the rebel movement as it sought to expel rivals from its strongholds and asserted itself as the sole legitimate rebel organization that merits allegiance (Lister 2015).

Cooperative Rebel Network: Data and Variables

We measure rebel cooperation (our dependent variable) in terms of claims of tactical JOINT OPERATIONS. The use of joint operations provides a more demanding test of ideological homophily than formal alliances because if groups prefer to cooperate with ideologically similar rebels at the tactical level, then they should be even more selective for the deeper, leadership-level collaboration required in strategic coalitions.

We began our data collection by tracking the operational claims of 44 major rebel groups using Arabic and English newspapers of record, US government informational briefs, and think tank reports. Since it was not possible to collect data on all the Syrian rebel groups, we limited our analysis to a medium *N* that had sufficient credible information to ensure data reliability. Although not ideal, expanding the analysis to less prominent groups risked sacrificing quality for quantity. Furthermore, by focusing on the primary rebel actors, we assumed as Krause (2017: 14) does that prominent players matter the most and that minor players are less likely to shape conflict trajectories.

We used automated text processing to find claims that contained "joint," "collaboration," "cooperation," or "support" and then hand coded each claim to verify it constituted a joint operation. We collected their claims of attacks—including both targets of attack and groups involved in joint operations. This data comes from US Government translations of insurgents' statements and operational claims, drawn from social media (Facebook, Twitter and YouTube) as well as various jihadist forums.

We used any claims of JOINT OPERATIONS from the 44 organizations to one another or to smaller groups to generate a network of some 220 Sunni Arab and Kurdish groups actively engaged in the conflict. This resulted in 696 joint operations and more than 930 ties between the 220 groups across the four years of the conflict spanned by our data (July 2012 to June 2015). The JOINT OPERATIONS network is a symmetric matrix whose elements are the total number of JOINT OPERATIONS claimed by either group in the dyad represented. If more than two groups were claimed to be involved in an attack, we gave each group a tie with each other group mentioned.

Network Description

The full network (Figure 1) shows some clear patterns of cooperation. The more prominent groups, such as ASIM, ANF, ISIL, and Al-Sham Legion (ASL), have separate retinues of small groups linked only to them. However, there is also cooperation among prominent groups. ASIM, the group with the most ties, cooperates with large FSA-affiliated groups such as the Ahfad al-Rasul Brigades (AARB) and the Al-Furqan Battalions (AFB) as well as the sectarian jihadists ANF and ISIL. There are Kurdish groups in our data, observed in the lower right corner and linked to the Sunni Arab militants by a single connection—YPG to Ar-Raqqa Revolutionaries Brigade (ARRB).

[Figure 1 about here]

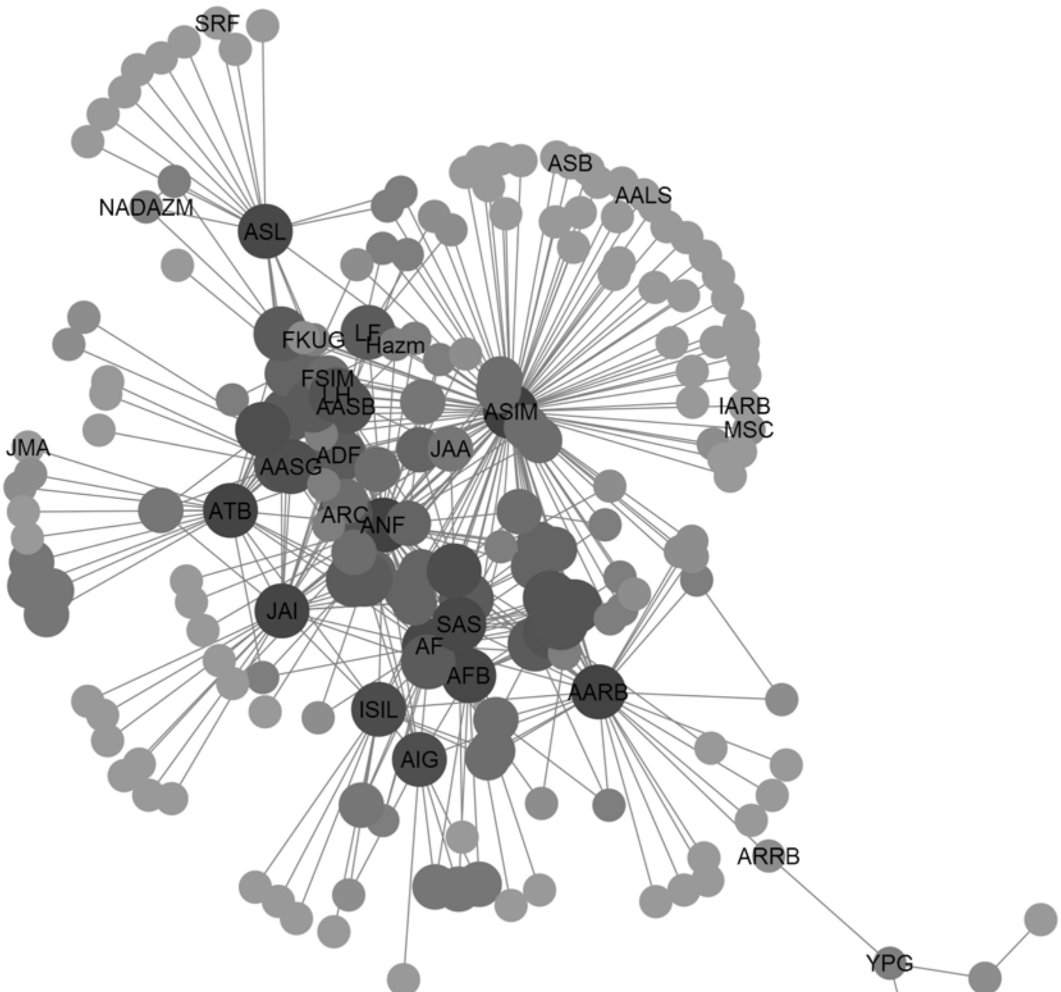

Figure 1. Network of Syrian militant JOINT OPERATIONS. Links indicate presence of one or more ties between groups. Circle size is proportional to node degree. The names of smaller groups have been removed from this graph to make it readable.

Measuring Ideology, Power, and State Sponsorship in the Network

We evaluated rebel groups for three ideological areas of relevance to the Syrian conflict. Sectarianism serves as our CONFLICT FRAME variable: groups with high sectarianism scores cast the conflict as Sunnis vs. Shiites/Alawites, whereas groups with low sectarianism scores have little or no anti-Shiite rhetoric. Salafism, which measures the extent to which groups ascribe to that puritanical strain of Sunni Islam, provides our IDEAL POLITY variable. The use of Salafism better resolves differences within various stripes of Islamists than a simple secularism vs. Islamism scale. Revisionism is used for the TERRITORIAL ASPIRATION component of ideology: groups with low scores seek to preserve Syria's territorial integrity, whereas a high score signifies a desire to abrogate it, in particular as do Caliphate-minded sectarian jihadists or Kurdish separatists.

Each axis of ideology is coded on a scale of 1-5, a range that allows us to capture the proximity or distance of groups on each component. We hand coded the ideology of the 44 Syrian rebel groups using manual coding from the groups' founding charters and other public declarations (see *Supplemental Material* for coding methodology and the rebel groups' ideological scores). We aggregate the scores of the three components of ideology into one AVERAGE IDEOLOGY score and check those results to make sure the variable we have constructed makes sense in light of Syria's factional divides (see the *Supplemental Material*).

Our methodology situates groups in ways that make sense in the context of the Syrian Civil War. We would expect groups in Syria to broadly fall into the following categories: *Secular nationalists, Salafist nationalists, secular Kurdish separatists, and sectarian jihadists* (see Table 2).[8]

[Table 2 about here]

Table 2: Ideological spectrum of Syria's militant factions

Dimensions of Ideology	Secular nationalist	Kurdish separatist	Salafist nationalist	Sectarian jihadist
Conflict Frame	Syrians vs. the Assad regime	Kurds vs. sectarian jihadists	Sunni Syrians vs. the Assad regime	Sunnis vs. Alawites/Shiites
Ideal Polity	Secular democratic Syria	Kurdish secular government	Islamic state	Islamic state
Territorial Aspiration	Unified Syria	Separate republic or autonomy	Unified Syria	Transnational Islamic Caliphate

We measure POWER in the network by group size. For each group, we collected as many estimates of size as were available (see *Supplemental Material*). We created a low-medium-high estimate for each group, when possible, and used the medium estimate in our analysis. As we could not locate size estimates for a few small groups, they were assigned a minimal value of 500 fighters.

Although group size is by no means a comprehensive measure of power, it is often used as such in statistical analysis (Akcinaroglu 2012; Christia 2012; Krause 2013/14). We make the assumption that groups that can mobilize more fighters than their competitors are also likely to have substantial financial resources to arm those fighters, pay them salaries, and support their families. Thus, we proceed with group size as a proxy for other elements of rebel power. We also use the Institute for the Study of War (ISW) "powerbrokers" measure to validate this variable and find that no group coded as a regional powerbroker by ISW is also a "small" rebel group in terms of number of fighters (Cafarella and Casagrande 2016). The smallest group in our data listed as a powerbroker is Nur al-Din al-Zinki Movement (NADAZM) with a size estimate of approximately 5,000 fighters.

Lastly, we assess the presence or absence of SHARED STATE SPONSORSHIP by drawing upon informed secondary sources that identify the primary sponsors of rebel groups (see the *Supplemental Material* for the complete list of sources). As for rebel group LOCATION, we used the operational claims of rebels to determine their primary areas of operation. Some groups operated locally, and were coded as such, while others had multiple branches. Groups that appeared in four or more governorates were coded as national, even though they may not have had presence in every Syrian region.

The Core Network

We coded covariate data for 44 rebel groups and tracked the collaborative relationships among them. Only 31 of those 44 groups participated in collaborative tactical relationships, so we proceed with 376 ties among these 31 groups (see Figure 2).

[Figure 2 about here]

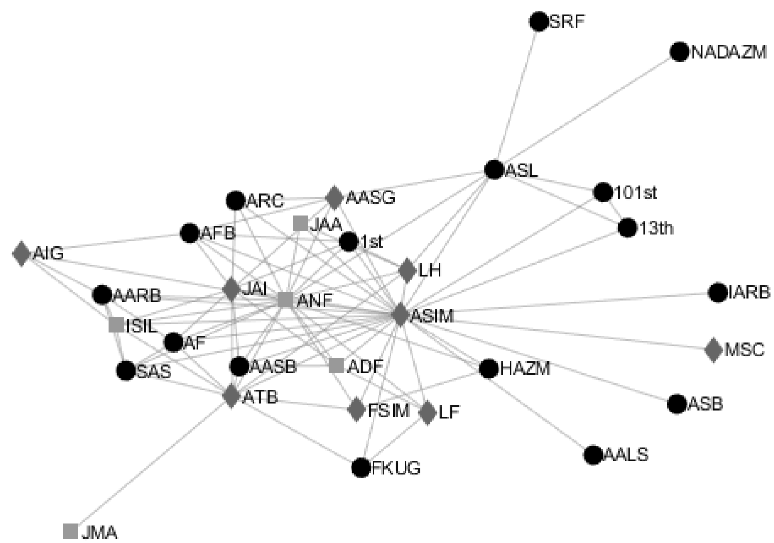

Figure 2. Diagram of the 31 core groups with collaborative ties within the network. A line between two groups indicates the presence of at least one joint operation between them. Node shapes denote secular nationalist (circles), Salafist nationalist (diamonds), and sectarian jihadist (squares) ideological classifications.

The *Supplemental Material* provides data on the core groups used in this analysis, as well as descriptive statistics. The core groups display significant variation in terms of group size, ideology and state sponsorship.

Network Regression Analysis

To evaluate our hypotheses relative to one another and while controlling for additional variables, we run an Additive and Multiplicative Effects (AME) regression models from Minhas, Hoff and Ward (2016a). The AME regression is an extension of the class of network inference methods known as latent space models, which seek to relate the tendency of nodes to form ties with each other to their proximity in an underlying space of latent variables (Cranmer, McClurg and Rolfe 2017). Rather than assuming independence of observations (as per de Finetti's theorem, which justifies the conditionally independent and identically distributed assumption in statistics), these models account for dependence between dyads (row and column means – additive effects) and higher order dependence in the network structure, such as stochastic equivalence (multiplicative effects – see *Supplemental Material* for model specifications and details). AME models have been used recently to analyze conflict and international relations data (see Dorff 2015; Dorff, Gallop and Minhas *forthcoming*, Minhas, Hoff and Ward 2016b).

Table 3 presents the AME regression results for difference in AVERAGE IDEOLOGY, difference in POWER, and shared STATE SPONSORSHIP separately (Models 1-3) and together with controls for their node-level values (for state sponsorship, a dummy variable corresponding to the presence or absence of a sponsor) and SHARED LOCATION (Model 4) and a control for our most prominent group, ASIM (Model 5). As the AME software does not yet cover Poisson or Negative Binomial Distributions, we follow standard practice for such cases and take the square root of the dependent variable to account for the progressively increasing residuals in our data as independent variable values increase. Results using the direct count value of the dependent variable and an ordinal model are displayed in *Supplemental Material.*

[Table 3 about here]

Table 3. Square root transformed dependent variable.

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.07 (0.11)	0.75 (0.00)	-0.00 (0.00)	-0.48 (0.28)	-0.41 (0.32)
State Sponsorship (Node)				0.07 (0.11)	0.09 (0.12)
Ave. Ideology (Node)				0.07 (0.04)	0.07 (0.04)
Power (Node)				0.01 (0.01)	0.00 (0.01)
ASIM (Node)					0.32 (0.28)
Ideol. Diff. (Dyad)	-0.04*** (0.01)			-0.05*** (0.00)	-0.05*** (0.01)
Power Diff. (Dyad)		-0.06* (0.04)		-0.01*** (0.00)	-0.01*** (0.00)
Shared St. Sponsor (Dyad)			0.06 (0.05)	-0.00 (0.05)	-0.10 (0.05)
Shared Location (Dyad)				0.17*** (0.04)	0.17*** (0.04)

Results of AME regression analysis. Dependent variable is square root of the count of collaborative ties. Standard Errors in parenthesis. *p <0.05; ** p<0.01; ***p<0.001

Tables 3 reveals strong support for the ideological alliances hypothesis (*H_1*): groups that are ideologically proximate cooperate with each other more so than the ones that are ideologically distant. The decrease in cooperation with increasing ideological difference is statistically significant regardless of the inclusion of other covariates and whether the square root transformation or raw counts is used.

We find some support for power symmetry (*H_2*): groups of similar strength tend to cooperate with each other more so than those that vary in their power capabilities. This finding, however, is inconsistent. It is always significant in the square root transformation but, in the direct tie count, only becomes significant for Model 4 (see *Supplemental Material*). The support for power symmetry rules out the opposite prediction of power asymmetry (*H_3*).

Ideology has a consistent, statistically significant value across models with a smaller degree of uncertainty than for the power difference (see *Supplemental Material*), and with a larger effect size than power in the majority of models. Also, the robustness checks below firmly support the ideological homophily hypothesis, but not power symmetry or asymmetry.

We do not find evidence in favor of the state-sponsored alliances hypothesis (*H_4*), which predicts that rebel dyads with shared sponsors should exhibit greater cooperation than dyads with distinct external sponsors. Perhaps this result is an artifact of the lack of weight in shared sponsorship; our data represents whether two groups were ever sponsored by the same state, with no weight given to how important a sponsor was for a particular group. More substantively, however, two plausible explanations may shed light on this finding. First, as long as rebel groups were generally cooperating with ideologically similar groups, which is what we find in the Syrian civil war, their state sponsors may not have cared if their clients were working with rebels that have different sponsors. However, had it been the case that Syrian groups were, generally speaking, cooperating with ideologically dissimilar factions, their state sponsors would have exerted pressure to break those cooperative ties.

A second possible explanation is that tactical cooperation in the form of joint operations is less visible to state sponsors than strategic mergers or formal coalitions. States may have overlooked their clients' tactical cooperation partners in the Syrian theater to achieve the broader objective of regime change. It is more likely that overlapping external sponsorship plays a greater role in facilitating or hindering strategic alliances that are much more formal and public than they do tactical cooperation. Therefore, our finding regarding state sponsorship at the tactical cooperation level does not preclude the importance of this variable in strategic alliances, which we do not explore in this study. [9]

Table 3 includes SHARED LOCATION as a control since it is possible that observed cooperation between ideologically similar groups is merely a surface manifestation of the underlying ideological homogeneity of groups who operate in the same area. Tactical joint operations, perforce, require rebels to operate in the same location and if these operating areas consisted only of groups with similar ideologies then ideological homophily would arise simply due to geographic proximity. One might argue that the homogeneity of a given geographic area with respect to its ethnic or religious composition would foster such ideological homogeneity. Alternatively, one might expect that social influence between proximate groups would result in ideological convergence. This argument, however, begs the question as to the epiphenomenal nature of ideology by assuming that it is easily malleable in the first place. In counterpoint, contact between ideologically distinct groups may readily result in their violent conflict rather than ideological convergence, an outcome amply demonstrated in the Syrian conflict.

Empirically, Table 3 supports SHARED LOCATION as being important to tactical collaboration as intuitively expected, but ideological homophily still remains significant. The above supposition that geographic proximity imposes ideological uniformity is at odds with the fact that the predominantly Sunni Arab composition of the rebel movement reflected ideological diversity—a diversity that existed in close geographic proximity. In Aleppo, for example, there were no less than 22 separate armed groups representing three distinct ideological strands: secular nationalism, Salafist nationalism, and sectarian jihadism. Similarly, in Idlib, there were 16 groups representing these three distinct ideologies. In fact, nine of the 10 Syrian governorates in our study had at least two ideologically divergent groups (see *Supplemental Material*).

Robustness Checks: Assortativity and Network Simulation

Two additional network methods are employed as robustness tests for ideology and power in the above analysis: (1) comparing the assortativity, a metric of network homophily or heterophily, of the observed network with the distribution obtained from a null model simulation; and (2) a simulation with homophily included to estimate the characteristic ideological or power scale over which cooperation is more likely (a heterophily simulation is used if power asymmetry is indicated). These two methods consider the three ideological components, AVERAGE IDEOLOGY, and POWER – all treated separately.

Assortativity is the standard metric for assessing whether tie formation is driven by similarity or dissimilarity with respect to a scalar variable (as we operationalize power and ideology). The assortativity α is the correlation of the variable values for the nodes at each end of a tie taken over all ties (see Network Simulation Appendix in the *Supplemental Material*). An α value of +1 signifies maximal homophily whereas -1 represents maximal heterophily. For statistical testing purposes, the assortativity cannot be treated as one would treat a standard correlation because ties are not taken to be independent.

We developed a simple simulation of the tie formation process that can be implemented using our empirical data. When ideology (or power) is not included, the simulation acts as a null model that can generate a distribution of assortativity values for calculating the statistical significance of the observed assortativity. Our simulation-based tests will decide that homophily (heterophily) is present when the difference between the empirical network assortativity and the mean of the null simulation over many runs is positive (negative) and statistically significant.

In the simulation, nodes form ties (i.e. groups conduct joint operations) probabilistically. Each iteration consists of the placement of a tie between nodes where the iterations proceed up to the total number of ties in the observed network. The simulation is constrained in that it seeks to reproduce the node degrees in the observed network. Each node can only receive a maximum number of ties equal to its observed degree (its number of joint operations).[10] The model essentially assumes that a group wishes to make its units available for a certain number of joint operations over a given time period. The more units available at a given moment, the more likely a group is to find a partner, which, at the dyad level, implies that the interaction probability between a pair of groups depends on the product of their available units.

There are three variants of the simulation (see Network Simulation Appendix in the *Supplemental Material*): (1) a null simulation in which, as described above, only the node degrees affect tie formation, not the node variables; (2) a homophily simulation in which the probability of tie formation between two nodes decreases as the distance between their variable (ideology or power) values gets larger; and (3) a heterophily simulation in which the probability of tie formation increases with increasing distance between them. For the homophily simulation, a parameter called the interaction length, *l*, sets the characteristic length scale so that, roughly speaking, nodes are significantly more likely to form ties within that length from each other than beyond that range. As the interaction length scale increases, the effect of homophily diminishes until the null model is effectively recovered. The heterophily simulation uses a different parameter, the suppression length *ls*, for which the probability of interaction is reduced for the region within the suppression length and much greater in the region outside it.

Robustness Findings

The statistical tests using ASSORTATIVITY are shown in Table 4 for the entire 2012-2015 period and by individual years (note that 2012 and 2015 are not full years of data). For the entire period, the observed ASSORTATIVITY values for ideology are all greater than the mean of the null distribution indicating homophily. CONFLICT FRAME and TERRITORIAL ASPIRATION are highly significant. Although IDEAL POLITY is not significant, the mean of all three components, AVERAGE IDEOLOGY, remains highly significant. Accordingly, the assortativity tests support the ideological homophily hypothesis (H_1); ideological clustering characterizes the network structure.[11] The results of the simulation, which models interactions driven by ideological homophily, are similar in that CONFLICT FRAME, TERRITORIAL ASPIRATION, and AVERAGE IDEOLOGY all display well-defined values of the interaction length *l* whose confidence intervals are less than the full range of the ideology scale, whereas IDEAL POLITY does not. The common value of l = 2.3 indicates that the zone in which cooperation is relatively likely is about half the length of the full ideology scale. Thus, groups in the middle of the spectrum can cooperate with both ends, but cooperation between the opposed extremes of the scale will be much less common.

For the network by year, the assortativity for AVERAGE IDEOLOGY is highly significant for 2013 and 2014, the two years with the greatest number of ties. The 2012 and 2015 networks, which are smaller, show no significance.

[Table 4 about here]

Table 4. Robustness Analysis Results

VARIABLE	ASSORTATIVITY α	α_{null}	σ_{null}	*p*	SIMULATION *l*/l_S	*CI*
2012-2015 (*N*=376)						
Conflict Frame	-0.032(+)***	-0.196	0.039	<.0001	2.3	(2.0,2.7)
Ideal Polity	-0.096(+)	-0.097	0.047	.97	--	--
Territorial Aspiration	-0.027(+)***	-0.238	0.042	<.0001	2.3	(2.0,2.8)
Average Ideology	0.017(+)***	-0.145	0.042	<.0001	2.3	(1.7,3.6)
Power	-0.235(–)	-0.172	0.048	.19	--	--
2012 (*N*=55)						
Average Ideology	-0.138 (–)	-0.134	0.108	.96	3.1	(1.9,5.4)
Power	-0.574 (–)	-0.382	0.110	.07	5900 (l_S)	(2500,10000)
2013 (*N*=136)						
Average Ideology	0.096 (+)**	-0.122	0.072	.003	1.8	(1.5,2.2)
Power	-0.31 (–)*	-0.152	0.078	.04	6300 (l_S)	(3000,11000)
2014 (*N*=119)						
Average Ideology	0.025 (+)***	-0.230	0.075	.0006	--	--
Power	-0.236 (–)	-0.176	0.081	.47	--	--
2015 (*N*=66)						
Average Ideology	-0.313 (–)	-0.242	0.104	.51	--	--
Power	-0.141 (+)	-0.146	0.107	.94	8900	(6800,14000)

N is the number of ties. For Assortativity: α is the assortativity of the observed network where + (-) indicates α greater (less) than α_{null} corresponding to homophily (heterophily); α_{null} and σ_{null} are respectively the mean and standard deviation of the assortativity in the null simulation taken over 10,000 runs; the p-value is the (two-tailed) fraction of runs exceeding $|\alpha - \alpha_{null}|$. For Simulation: *l* is the mean interaction length (suppression length l_S where indicated) and CI is the 95% confidence interval (blank entries signify the absence of a clear minimum); 1,000 runs at each point were used to generate 1000 resamples of size 50 with replacement and then the minimum *1* (or l_S) for each resample was found. ***p < .001, **p < .01, *p < .05.

The substantive effect of ideological homophily can be assessed by running the homophily simulation with the estimated interaction length from Table 4 and examining how the number of ties for a dyad depends upon their ideological separation. For example, the interaction length for AVERAGE IDEOLOGY in the 2012-2015 network is 2.3. To enable a more generic assessment not contingent upon the specific Syria configuration, we simulated a network with a uniform degree distribution and uniform ideology distribution and found that increasing the distance between two nodes initially collocated at the middle of the ideology range to successive distances of (1, 2, 3, 4) units decreased the probability of tie formation between them by (5.4, 20.5, 39.2, 56.6) percent relative to the probability at zero distance (see *Supplemental Material*). Although the probability of tie formation depends nonlinearly on the ideological distance, averaging the above changes yields a 14.15 percent drop in probability per unit of distance. This value is consistent with the 13.4 percent decrease per one unit shift in ideological difference found by taking the beta value from AME Model 5 and running a simple linear prediction function while varying the values of ideological difference.

Turning to POWER, the observed assortativity for the full period is less than the mean of the null simulation, indicating a tendency toward heterophily, but is not significant. Given this tendency, the heterophily simulation was performed but no well-defined suppression length was found. The assortativity tests for the first three individual years also indicate a heterophily tendency, but which only rises to significance for 2013. On balance, the assortativity tests do not support either power symmetry or asymmetry.

When comparing the findings of the AME and robustness analyses, both show statistically significant results for ideological homophily, thus providing comprehensive support for ideology as an important factor in determining rebel collaboration. However, the two analyses differ over power. The AME analysis indicates statistically significant power symmetry while the assortativity analysis points toward power asymmetry, albeit not significant. The reason for this disparity may arise from a nonlinear relationship between tie formation and group power differentials. Considering the difference between the observed number of ties in a dyad and that expected based solely from the group degrees (Equation 1 in the *Supplemental Material*), the distribution of this quantity as calculated from the network (no simulation involved) shows an inverted U-shape as a function of power difference rather than a monotonic decrease (increase) as would be fully consistent with a power symmetry (asymmetry) dynamic (see *Supplemental Material*): the observed-expected ties difference is negative at both low (< 5,000) and high (> 20,000) power differences and positive for intermediate ones. The tie suppression at low power difference is consistent with power asymmetry whereas that at high power difference is consistent with power symmetry. In contrast, a similar plot for ideology shows a greater than expected number of ties for low ideology differences (< 2) and smaller than expected for higher ideology differences, and so is clearly consistent with an overall homophily effect.

Finally, we address the concern that the finding of ideological cooperation may be an artifact of our limited sample of only 44 groups out of hundreds in the Syrian conflict. As the set of omitted groups is almost entirely, if not completely, composed of relatively weak groups, this concern amounts to the possibility that the weaker groups in our sample are unrepresentative of the broader universe of weak groups in Syria. Since the strong groups are ultimately of greatest importance, we test for ideological homophily between them. Indeed, considering the network of joint operations between groups of size at least 5,000 over the full time period, the assortativity test finds homophily for AVERAGE IDEOLOGY to be significant ($p<.0001$). Therefore, ideological similarity helps drive cooperation between the groups whose behavior is most consequential to the conflict. Additionally, a t-test reveals no significant difference between the AVERAGE IDEOLOGY means of the strong and weak (< 5,000) groups in our sample so there is no basis to believe that our sample of weak groups is unrepresentative.

Discussion and Conclusion

Rebel cooperation is a common occurrence in civil conflicts. In this study, we wanted to know with whom rebels cooperate in the context of fragmented conflicts that feature a diversity of ideological actors, variation in group-level power capabilities, and a plethora of state sponsors. Theoretically, we proposed three components of militant ideology and argued how each can facilitate cooperation. Conflict framing promotes shared understandings about ingroups and outgroups, thereby easing potential dissension about permissible allies and targets. A similar conception of the ideal polity encourages groups to work cooperatively toward achieving compatible visions of the post-conflict political order. Territorial aspiration impacts fundamental questions such as whether or not rebels seek to break up the national state or maintain its integrity, which are incompatible goals that dampen cooperation. We employed an innovative network-analytic methodology, constructing a militant tactical cooperation network from claims of joint operations and relating its structure to ideology, power, and state sponsorship in the Syrian civil war.

Substantively, our central finding is that ideology is an important determinant of alliance composition in the Syrian civil war. Groups that were ideologically similar cooperated more frequently than those who were ideologically dissimilar: according to our models, a one unit increase in ideological distance corresponds to about a 14% decrease in the likelihood of rebel tactical cooperation. Syrian groups in the middle of the ideological spectrum were willing to cooperate with groups at the end of the spectrum, but groups at the end of the ideological spectrum were less willing to cooperate with each other. No clear finding concerning power emerged as one analysis supported power symmetry while the other supported neither symmetry or asymmetry. We found no evidence that having a common state sponsor encouraged cooperation.

Our ideology hypothesis and results may elicit an endogeneity objection. It could be asserted that stable interrebel relationships motivated by power form first and then groups adjust their ideologies accordingly. In this scenario, power drives the ideological preferences exhibited in alliance composition. This challenge assumes that militant groups arise as ideological blank slates, contrary to the fact that the founders of such groups often have strong ideological orientations. Many of the individuals who formed Syria's major Islamist rebel groups were actually in jail at the start of the revolution due to their prior Islamist activism and then subsequently released (Lister 2015, 53-55; Baczko, Dorronsoro, Quesnay 2018, 184).[12] In addition, ideological charters, an important element of our coding, are typically issued by groups shortly after their formation. Their ideological statements, therefore, are biased toward a time before these groups have formed cooperative ties with other rebels, and so evidence of homophily in the network reflects selection of similar others.

Another endogeneity concern is that we treat ideologies as fixed, but conflict processes are likely to change the ideological preferences of rebel factions over time. We treat the question of ideological change as an empirical one and our operationalization of ideology using the three components can help track that change. We suspect that ideologies change over time but do so gradually. A common process in social networks is increasing homogeneity in network ties because of the selection of similar partners and the reinforcing effect of social influence of those partners in maintaining that similarity. Ideological shifts, therefore, will typically be limited and evolutionary, a process that allows for stable patterns to emerge between ideology and cooperative network structure. This view of incremental ideological change is supported by our findings of ideological homophily for the full 2012-2015 network and individually for 2013 and 2014 using ideology scores biased toward earlier times in group histories.

Another concern relates to the interaction between power balancing and ideological considerations. These broad concepts are not alternative and incompatible explanations of alliance composition. Rebels may form balanced coalitions, each of which consists of ideologically similar groups. As rebels face a greater (lesser) threat from the state, they may become less (more) ideologically selective about their allies. However, ideology may also act as a barrier to alliance formation even when the distribution of power is so adverse that it would seem to demand rebel unification. In Syria, the tide turned dramatically against the rebels after Russia's direct military involvement on the side of the regime in late 2015 and after the fall of Aleppo in late 2016. Yet, the rebels did not ally across ideological lines but remained bitterly divided (Collins 2017; Perry and Al-Khalidi 2017).

Our empirical analysis of a single case study limits our ability to generalize beyond the Syrian conflict. Although not entirely unique, the Syrian civil war is characterized by severe levels of movement fragmentation, a wide spectrum of ideologies, and a perplexing array of external interventions by state and sub-state actors. Therefore, it may not be entirely representative of other civil wars where rebel groups are fewer in number, non-ideological identities prevail (such as in ethnic or resource-based conflicts), or where international interference is limited in scope. Our findings regarding the robustness of ideological homophily in Syrian militant alliances should be thoroughly investigated in other conflicts to have confidence in its generalizability. Our Syrian study, however, highlights the need to consider seriously the role of ideology in rebel alliances and offers a template for researching civil conflicts that exhibit similar patterns of intense fragmentation, ideological polarization, and tactical alliances, such as those ongoing in Ukraine, Iraq, Afghanistan, Pakistan, Libya, Sudan, and Yemen.

Acknowledgments

We thank the two anonymous reviewers. We also thank Cassy Dorff, Peter Hoff, Peter Krause, Lee Seymour, Sarah Parkinson, Jon Mercer, Theo MacLauchlan and Mike Weintraub or their helpful comments.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This research was supported by the Defense Threat Reduction Agency and the Office of Naval Research under grants HDTRA1-10-1-0075, N00014-15-1-2549, and N00014-16-1-2919.

Supplemental Material

Supplemental material for this article is available online.

References

Akcinaroglu, Seden. 2012. “Rebel Interdependencies and Civil War Outcomes.” *Journal of Conflict Resolution* 56 (5): 879-903.

Asal, Victor, and R. Karl Rethemeyer. 2008. “The Nature of the Beast: Organizational Structures and the Lethality of Terrorist Attacks.” *Journal of Politics* 70 (2): 437-449.

Baczko, Adam, Gilles Dorronsoro, and Arthur Quesnay. 2018. *Civil War in Syria: Mobilization and Competing Social Orders*. New York: Cambridge University Press.

Bakke, Kristin M. 2014 “Help Wanted? The Mixed Record of Foreign Fighters in Domestic Insurgencies.” *International Security* 38 (4): 150-187.

Bakke, Kristin M., Kathleen Gallagher Cunningham, and Lee J. M. Seymour. 2012. “A Plague of Initials: Fragmentation, Cohesion, and Infighting in Civil Wars.” *Perspectives on Politics* 10 (2): 265-283.

Balcells, Laia. 2017. *Rivalry and Revenge: The Politics of Violence during Civil War*. New York: Cambridge University Press.

Balliet, Daniel, Joshua M. Tybur, Junhui Wu, Christian Antonellis, and Paul A. M. Van Lange. 2016. “Political Ideology, Trust, and Cooperation: In-group Favoritism among Republicans and Democrats during a US National Election.” *Journal of Conflict Resolution* (July): 1-22.

Bapat, Navin A., and Kanisha D. Bond. 2012. “Alliances between Militant Groups.” *British Journal of Political Science* 42 (4): 793-824.

Boutyline, Andrei, and Robb Willer. 2017. “The Social Structure of Political Echo Chambers: Variation in Ideological Homophily in Online Networks.” *Political Psychology* 38 (3): 551-569.

Cafarella, Jennifer, and Genevieve Casagrande. 2016. *Syrian Armed Opposition Powerbrokers*. http://www.understandingwar.org/sites/default/files/Syrian%20Armed%20Opposition%20Powerbrokers_0_0.pdf (accessed December 2, 2017).

Cranmer, Skyler J., Leifeld, Philip, McClurg, Scott D., & Rolfe, Meredith. (2017). Navigating the Range of Statistical Tools for Inferential Network Analysis. *American Journal of Political Science, 61*(1), 237-251.

Christia, Fotini. 2012. *Alliance Formation in Civil Wars*. New York: Cambridge University Press.

Collins, Dylan. 2017. "Idlib's Rebel Split: A Crossroads for Syrian Opposition." *Al-Jazeera* (February 02).

Corbetta, Renato. 2013. "Cooperative and Antagonistic Networks: Multidimensional Affinity and Intervention in Ongoing Conflicts, 1946-2001." *International Studies Quarterly* 57 (2): 370-384.

Costalli, Stefano, and Andrea Ruggeri. 2015. "Indignation, Ideologies, and Armed Mobilization: Civil War in Italy, 1943-45." *International Security* 40 (2): 119-157.

Cunningham, Kathleen Gallagher. 2013. "Actor Fragmentation and Civil War Bargaining: How Internal Divisions Generate Civil Conflict." *American Journal of Political Science* 57 (3): 659-672.

Cunningham, Kathleen Gallagher, Kristin M. Bakke, and Lee J. M. Seymour. 2012. "Shirts Today, Skins Tomorrow: Dual Contests and the Effects of Fragmentation in Self-Determination." *Journal of Conflict Resolution* 56 (1): 67-93.

DellaPosta, Daniel, Yongren Shi, and Michael Macy. 2015. "Why Do Liberals Drink Lattes?" *The American Journal of Sociology* 120 (5): 1473-1511.

Dorussen, Han, Erik A. Gartzke, and Oliver Westerwinter. 2016. “Networked International Politics: Complex Interdependence and the Diffusion of Conflict and Peace.” *Journal of Peace Research* 53 (3): 283-91.

Dorff, Cassy L. 2015. “Civilian Autonomy and Resilience in the Midst of Armed Conflict.” *Ph.D. Dissertation.*

Dorff, Cassy L., Max Gallop and Shahryar Minhas. *Forthcoming*. “Networks of Violence: Predicting Conflict in Nigeria.” *Journal of Politics.*

Driscoll, Jesse. 2012. “Commitment Problems or Bidding Wars? Rebel Fragmentation as Peace Building.” *Journal of Conflict Resolution* 56 (1): 118-149.

Fearon, James D. and David D. Laitin. 2000. “Violence and the Social Construction of Ethnic Identity.” *International Organization* 54 (4): 845-877.

Gerber, Elisabeth R., Adam Douglas Henry, and Mark Lubell. 2013. “Political Homophily and Collaboration in Regional Planning Networks.” *American Journal of Political Science* 57 (3): 598-610.

Griffiths, Ryan D. 2016. *Age of Secession: The International and Domestic Determinants of State Birth*. New York: Cambridge University Press.

Gurr, Ted R. 1970. *Why Men Rebel*. Princeton, NJ: Princeton University Press.

Gutiérrez Sanín, Francisco, and Elisabeth Jean Wood. 2014. “Ideology in Civil War: Instrumental Adoption and Beyond.” *Journal of Peace Research* 51 (2): 213-226.

Hafner-Burton, Emilie M., Miles Kahler, and Alexander H. Montgomery. 2009. “Network Analysis for International Relations.” *International Organization* 63 (3): 559-92.

Hoff, Peter D. 2015. *Dyadic data analysis with AMEN*. https://arxiv.org/abs/1506.08237 (accessed June 1, 2017).

Hoover Green, Amelia. 2016. “The commander’s dilemma: Creating and controlling armed group violence.” *Journal of Peace Research* 53 (5): 619-632.

International Crisis Group. 2012. *Tentative Jihad: Syria’s Fundamentalist Opposition*. https://d2071andvip0wj.cloudfront.net/131-tentative-jihad-syria-s-fundamentalist-opposition.pdf (accessed March 24, 2017).

International Crisis Group. 2014. *Flight of Icarus? The PYD’s Precarious Rise in Syria*. https://www.crisisgroup.org/middle-east-north-africa/eastern-mediterranean/syria/flight-icarus-pyd-s-precarious-rise-syria (accessed March 24, 2017).

Jones, Seth G. 2017. *Waging Insurgent Warfare: Lessons from the Vietcong to the Islamic State*. Oxford University Press.

Kim, Nam Kyu. 2018. “Revolutionary Leaders and Mass Killing.” *Journal of Conflict Resolution* 62 (2): 289-317.

Krause, Peter. 2013/14. “The Structure of Success: How the Internal Distribution of Power Drives Armed Group Behavior and National Movement Effectiveness.” *International Security* 38 (3): 72-116.

Krause, Peter. 2017. *Rebel Power: Why National Movements Compete, Fight, and Win*. Ithaca, NY: Cornell University Press.

Lai, Brian, and Dan Reiter. 2000. “Democracy, Political Similarity, and International Alliances, 1816-1992.” *Journal of Conflict Resolution* 44 (2): 203-227.

Lichbach, Mark I. 1995. *The Rebel’s Dilemma*. Ann Arbor, MI: Michigan University Press.

Lister, Charles R. 2015. *The Syrian Jihad: Al-Qaeda, the Islamic State, and the Evolution of an Insurgency*. New York: Oxford University Press.

Maoz, Zeev. 2012. “Preferential Attachment, Homophily, and the Structure of International Networks, 1816-2003.” *Conflict Management and Peace Science* 29 (3): 341-369.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. “Birds of a Feather: Homophily in Social Networks.” *Annual Review of Sociology* 27: 415-444.

Metternich, Nils W., Cassy Dorff, Max Gallop, Simon Weschle, and Michael D. Ward. 2013. “Antigovernment Networks in Civil Conflicts: How Network Structures Affect Conflictual Behavior.” *American Journal of Political Science* 57(4): 892-911.

Minhas, Shahryar, Peter D. Hoff and Michael D. Ward. “Inferential Approaches for Network Analysis : AMEN for Latent Factor Models I.” (2016a).

Minhas, Shahryar, Peter D. Hoff, and Michael D. Ward. “A new approach to analyzing coevolving longitudinal networks in international relations.” *Journal of Peace Research* 53, no. 3 (2016b): 491-505.

Morrow, James D. 1991. “Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances.” *American Journal of Political Science* 35(4): 904-933.

Oppenheim, Benjamin, Abbey Steele, Juan Vargas, and Michael Weintraub. 2015. “True Believers, Deserters, and Traitors: Who Leaves Insurgent Groups and Why.” *Journal of Conflict Resolution* 59 (5): 794-823.

Perry, Tom and Suleiman Al-Khalidi. 2017. “Syrian Rebels Bitterly Divided Before New Peace Talks.” *Reuters* (January 20).

Phillips, Christopher. 2016. *The Battle for Syria: International Rivalry in the New Middle East*. New Haven: Yale University Press.

Popovic, Milos. 2018. “Inter-Rebel Alliances in the Shadow of Foreign Sponsors.” *International Interactions* 44 (4): 749-776.

Rudloff, Peter, and Michael G. Findley. 2016. "The Downstream Effects of Combatant Fragmentation on Civil War Recurrence." *Journal of Peace Research* 51 (1): 19-32.

Salehyan, Idean. 2010. "The Delegation of War to Rebel Organizations." *Journal of Conflict Resolution* 54 (3): 493-515.

Salehyan, Idean, David Siroky and Reed M. Wood. 2014. "External Rebel Sponsorship and Civilian Abuse: A Principal-agent Analysis of Wartime Atrocities." *International Organization* 68 (3): 633-661.

Seymour, Lee, Kristin M Bakke, and Kathleen Gallagher Cunningham. 2016. "E Pluribus Unum, Ex Uno Plures: Competition, Violence, and Fragmentation in Ethnopolitical Movements." *Journal of Peace Research* 53(1): 3-18.

Shesterinina, Anastasia. 2016. "Collective Threat Framing and Mobilization in Civil War," *American Political Science Review* 110 (3): 411-427.

Staniland, Paul. 2014. *Networks of Rebellion: Explaining Insurgent Cohesion and Collapse*. Ithaca: Cornell University Press.

Staniland, Paul. 2015. "Militias, Ideology, and the State." *Journal of Conflict Resolution* 59 (5): 770-793.

Tamm, Henning. 2016. "Rebel Leaders, Internal Rivals, and External Resources: How State Sponsors Affect Insurgent Cohesion." *International Studies Quarterly* 60 (4): 599–610.

Ugarriza, Juan E., and Matthew J. Craig. 2012. "The Relevance of Ideology to Contemporary Armed Conflicts: A Quantitative Analysis of Former Combatants in Colombia." *Journal of Conflict Resolution* 57 (3): 445-477.

Walter, Barbara F. 2017. "The Extremist's Advantage in Civil Wars." *International Security* 42 (2): 7-39.

Ward, Michael D., Katherine Stovel, and Audrey Sacks. 2011. "Network Analysis and Political Science." *Annual Review of Political Science* 14: 245-264.

Werner, Suzanne, and Douglas Lemke. 1997 "Opposites Do Not Attract: The Impact of Domestic Institutions, Power, and Prior Commitments on Alignment Choices." *International Studies Quarterly* 41 (3): 529-546

Wood, Reed M., and Jacob D. Kathman. 2015. "Competing for the Crown: Interrebel Competition and Civilian Targeting in Civil War." *Political Research Quarterly* 68 (1): 167-179.

Zech, Steven T., and Michael Gabbay. 2016. "Social Network Analysis in the Study of Terrorism and Insurgency: From Organization to Politics." *International Studies Review* 18 (2): 214-243.

Notes

[1] Ideological diversity emerges for a number of reasons. First, some conflicts are ideological at their core, such as the ones featuring communist and fascists in the twentieth century (e.g. Spanish civil war 1936-1939), or fundamentalists and secularists today. Second, prewar political mobilization based on ideological cleavages can extend into the civil war, shaping dynamics of rebel cohesion (Staniland, 2014) and violence between ideological rivals (Balcells, 2017). Third, the entry of transnational ideological actors such as al-Qaeda or Hezbollah can create ideological polarization, forcing rebel groups and their communities to take sides based on sharp ideological divides (Bakke, 2014).

[2] Fearon and Laitin (2000: 857) refer to this process as constructing “antagonistic identities;” Asal and Rethemeyer (2008: 438) refer to it as “othering;” and Shesterinina (2016, 417) terms it “collective threat framing.”

[3] Territorial aspirations have been at the root of many secessionist civil conflicts, resulting in 131 sovereign states coming into existence since 1945, “a threefold increase in 70 years” (Griffiths 2016, 1).

[4] Between states, Morrow (1991, 921-923) argues that capability aggregation drives symmetric alliances.

[5] This assumption is supported by Lichbach’s (1995: 19) observation that in rebel coalitions, the largest and richest organizations tend to pay a disproportionate cost for maintaining an alliance.

[6] In international relations, Morrow (1991, 921-923) finds alliances are more frequent between powerful and weak states than between those of comparable power.

[7] According to Jones (2017: 136), of 181 insurgencies between 1946 and 2015, 82% involved outside support.

[8] For similar categorization of rebel factions, see Cafarella and Casagrande (2016, 9) and Phillips (2016, 131-134).

[9] Additional model specifications, including bivariate relationships and AME diagnostic plots, are available in the *Supplemental Material*.

[10] It is not always possible to reproduce the degrees exactly, but the differences are typically small.

[11] A network visualization showing how groups cluster by ideology is included in the *Supplemental Material.*

[12] Ahrar al-Sham Islamic Movement and Jaysh al-Islam are two such groups. It is likely that the Assad regime released these leaders in a cynical ploy to affirm its narrative that the opposition consisted of jihadist terrorists, a strategy that implies that the regime, at least, believed that these men would act on their ideological predilections. Moreover, veterans of earlier jihads formed two other prominent factions, Al-Nusrah Front and the Islamic State, which suggests deep ideological commitments over time.

Appendix 3

Integrating Computational Modeling and Experiments: Toward a More Unified Theory of Social Influence

Integrating Computational Modeling and Experiment Toward a More Unified Theory of Social Influence

Michael Gabbay

Abstract

While computational models of social influence dynamics have been developed in a diverse array of disciplines, the proliferation of models and simulations has far outpaced their empirical support. Most recent modeling work has been oriented toward large networks and empirical contact has been mostly confined to showing that models display behaviors superficially similar to observed ones. In contrast, laboratory experiments on social influence and group decision making use small groups but typically set their sights on the evaluation of qualitative hypotheses thereby limiting their ability to inform modeling and simulation. This chapter proposes a research agenda that emphasizes the close integration of laboratory and computational approaches to the study of social influence, aimed at the quantitative testing of model predictions on experimental data. Recent work on group polarization that illustrates this approach is described. The integration of experiments and modeling will both advance the basic science of social influence and enable the development of more solidly grounded computational simulations for real-world applications.

INTRODUCTION

Social influence is a central element of many behavioral areas, such as public opinion change, radicalization, and group decision making — all of concern to public policy. It affects the process by which people form attitudes toward their governments, other population groups, or external actors. Group decision making applications span political, military, economic, and legal domains. If computational social science is to aid in understanding, anticipating, and shaping such real-world contexts, then the development of accurate and broadly applicable models of social influence is essential. This chapter proposes that a deliberate and concerted integration of experimental investigation and computational modeling is needed to develop these models, an effort that will also advance the fundamental knowledge of social influence dynamics.

Within social psychology, social influence has largely been studied via the experimental testing of discrete theoretical hypotheses that express how a dependent variable responds to a change in an independent variable in qualitative language (e.g., increases, decreases, inverted U-shape). Quantitative reasoning is often employed in the rationale for theoretical propositions and, to a much lesser extent, formal or computational models are also used to motivate them. While this process has been successful in revealing and extensively probing individual phenomena, it has been less effective at synthesizing and reconciling concurrent and competing processes. Such synthesis would better inform the development of computational models as to the relative strengths of different processes and their interaction. It is especially important if one seeks to apply a model to anticipate the behavior of a particular group of interest. For instance, the group polarization effect might predict that a group will pursue an extreme policy whereas majority influence points toward a moderate policy. Consequently, guidance as to how those two processes play out together is necessary in order to model the group's overall behavior. That guidance, of course, would need to be context dependent. As will be seen, the failure of group polarization theory to be integrated with broader social influence phenomena leads group polarization theory to predict that every group of like-minded members to become more extreme, regardless of their initial opinion distribution. This is not so for the frame-induced polarization theory to be described below that does integrate group polarization with majority influence and consensus pressure: the theory can explain a systematic tendency for like-minded groups to become more extreme while being able to predict that individual groups will not.

The difficulty of synthesis in traditional social influence research has not, however, deterred a surge of modeling research across a range of disciplines. This activity has not been an unalloyed good for computational social science as much of this work has proceeded with little regard for empirical support. Sizable and divergent streams of research have arisen around particular modeling approaches with murky domains of validity. This proliferation casts doubt upon the empirical relevance of the associated behavioral findings and complicates model selection and evaluation for applications.

The integrated approach advocated here calls for experiments designed with the explicit purpose of quantitatively testing computational models against data. It will help restore the balance between modeling and experimental validation. The development of computational models in conjunction with experiment will force researchers to reckon more intently on combining concurrent effects in order to make quantitative predictions. That is, more unified theory will have effects be caused by multiple factors that earlier work associates with separate hypotheses. A greater orientation of experiments toward testing models rather than seeking new effects will encourage replication efforts and so place empirical findings on more solid ground. The increased focus on synthesis and the inevitable failures of previously successful models as they are tested in new regimes will spur theoretical innovation as well. Eventually, this integrated modeling-experiment approach will lead to convergence upon a set of social influence models that have substantial experimental support and so can be confidently extended to larger scale systems or included within more complex simulations of particular application contexts.

This chapter first presents a brief survey of social influence and related group decision making research, along with a discussion of how standard hypothesis testing and also quantitative modeling have been employed. The second section provides an overview of opinion network modeling. Next, the quantitative testing of computational models on experimental data is illustrated using recent work on group polarization conducted by Gabbay, Kelly, Reedy and Gastil (2017), which also shows how the modeling goal of synthesizing concurrent effects can lead to new and more unified theory. The fourth section then sketches the envisioned integration of modeling and experiment and discusses its potential benefits.

SOCIAL INFLUENCE RESEARCH

The term social influence is broadly applicable to both attitudinal or behavioral effects of human interactions. We focus upon research involving attitudes, opinions, and judgments here, mental constructs that often guide behavior. Classic social influence phenomena include: conformity, the tendency and pressures toward consensus in groups; majority and minority influence, the ability of majorities and minorities respectively to sway group opinions; and group polarization, the tendency of discussion among like-minded people to make positions more extreme. Two primary types of influence routes have often been invoked as explanations of such behaviors, normative and informational. Normative influence refers to the operation of group and broader social norms in setting expectations as to appropriate opinions and the value of consensus. Informational influence is the acceptance of information from others as evidence about the reality of the subject under consideration. Informational influence typically involves the alignment of one's private and publicly-expressed judgments whereas public agreement need not imply private acceptance under normative influence.

The body of social influence research above has been established through a process of hypothesis testing via laboratory experiments. Groundbreaking studies appeared in the 1950s and 60s. (see Eagly and Chaiken (1993), Ch. 13). In a classic experiment by Asch, a substantial proportion of subjects suspended the clear evidence of their senses when faced with a majority of experimental confederates who stated that a clearly shorter line was longer, thereby demonstrating the power of conformity to induce public compliance. Inverting the direction of influence as the process under investigation, Moscovici and collaborators found in a color discrimination task that minorities who advocated consistent positions were more effective in swaying subjects than inconsistent minorities, leading to a theory that majorities primarily exert normative influence while minority influence occurs mostly via the informational route. Group polarization, the tendency of discussion among like-minded individuals to lead to more extreme opinions, is another element in the social influence canon. It has both informational and normative influence explanations and will be discussed in detail below.

Research on group decision making in contexts that allow for interpersonal persuasion also involves social influence. One strain of decision making research considers the performance of groups in comparison with individuals (Kerr & Scott, 2004). For example, the wisdom-of-the-crowds hypothesis holds that simply aggregating individual judgments over many individuals yields greater accuracy than the judgments of individual experts under the assumption that the members of the pooled population make independent judgments whose uncorrelated errors cancel. Arguments have been made that social influence can either impair this performance by inducing correlated errors or improve it when greater individual confidence tends to be associated with greater accuracy (Becker, Brackbill, & Centola, 2017).

The vast majority of experiments on social influence and decision making have been aimed at testing discrete qualitative hypotheses. A hypothesis is proposed concerning the direction of an effect on the dependent variable, increase or decrease, due to the variation of an independent variable, which is then tested statistically. If the amount of change in the dependent variable is in the theorized direction and improbably attributed to the null hypothesis of no effect, then the proposed hypothesis is said to be supported by the data.

Within the social influence and group decision-making domain, research affiliated with the literature on social decision schemes (SDS), which are essentially mathematical rules for combining group member initial preferences into a final decision, has most consistently pursued a model-based quantitative approach. Although initially concerned with juries and binary (innocent/guilty) decisions, the SDS program grew to include decisions concerning quantitative judgments such as monetary awards and budgets (Hinsz, 1999). While it has been very successful with respect to its original jury concentration (Devine, 2012), a shortcoming of the SDS program is that its focus on testing an array of aggregation rules has come at the expense of deeper theoretical and model development with respect to a specific opinion change process. This absence of a theoretical impetus inhibits generalization of the results to broader contexts — for example, when subgroup and social network structure is important or for general opinion dynamics in populations not associated with a focal decision point.

OPINION NETWORK MODELING

Although the experimental study of social influence has been conducted by social psychologists and, to a much lesser degree, in other social sciences such as political science, sociology, and economics, the modeling of social influence dynamics has extended beyond social science to fields including physics, applied mathematics, computer science, and electrical engineering (Castellano, Fortunato, & Loreto, 2009; Crandall, Cosley, Huttenlocher, Kleinberg, & Suri, 2008; Proskurnikov & Tempo, 2017). While the primary goal of opinion network models is to predict final opinions from initial ones, the models typically describe a process that occurs over time. This section briefly discusses approaches to opinion network modeling and their empirical application.

Many models of opinion change have been developed in the fields noted above and beyond, involving a great diversity in methodological and substantive choices. One major methodological division involves whether outcomes are produced deterministically or stochastically. A fundamental substantive division involves the way opinions are mathematically represented. A binary representation is clearly applicable to situations that ultimately involve a decision over two alternatives such as a political election. Alternatively, a continuous representation can account for gradations of opinion on an issue or for decisions involving either explicit numerical quantities such as budgets or that can be approximated as a spectrum of options ordered along some dimension (e.g., the extent of an escalatory military response). Binary (or discrete) opinion models tend to have stochastic interactions; continuous opinion models usually (but need not) employ deterministic interactions. The choice of opinion representation also constrains the basic process that governs how opinions change when nodes (a term for individuals within a network) interact. A binary opinion must either remain the same or flip when a node interacts with other nodes. In the voter model, a dyadic interaction is assumed whereby a node adopts the opinion held by a network neighbor selected at random whereas the majority rule model proceeds by selecting subgroups at random with all member nodes adopting the majority opinion in the subgroup (Castellano et al., 2009). Continuous opinion models, on the other hand, allow for incremental shifts in opinion where the amount of change is a function of the distance between node opinions. The DeGroot and Friedkin-Johnsen models, as well as the consensus protocol (popular in the engineering literature on control), use a linear dependence in which the shift is proportional to the opinion difference (DeGroot, 1974; Noah E. Friedkin & Johnsen, 2011; Olfati-Saber, Fax, & Murray, 2007). Bounded confidence models assume a hard opinion difference threshold, within which nodes interact linearly, but beyond which interaction produces no change (Lorenz, 2007). The nonlinear model of Gabbay (2007) uses a soft threshold so that, rather than vanishing completely, the interaction decays smoothly with distance.

The vast majority of papers on opinion network models make no contact with empirical data. They start with a model, reasoned to be plausible (sometimes on the basis of social psychology research but sometimes on an appeal to common sense) and then generate simulation results, often in combination with mathematical analysis, on phenomena such as how the time to reach consensus scales with system size, the conditions conducive to the formation of camps of rival opinions, or the ability of extremists or influential individuals to shift opinions. Usually, the focus is on large systems taken to be representative of population-scale behavior. Consequently, such empirical connections as are reported are usually on the level of noting that model-generated curves exhibit qualitatively similar shapes to relationships observed in naturally-occurring data from large population systems (Crandall et al., 2008; Düring, Markowich, Pietschmann, & Wolfram, 2009; Török et al., 2013). However, some models have been shown to quantitatively reproduce empirical relationships such as the distribution of votes in proportional elections (Burghardt, Rand, & Girvan, 2016; Fortunato & Castellano, 2007).

Application of opinion network models to laboratory experiments remains mostly confined to testing models developed within traditional fields of human behavioral research rather than from the physical sciences and engineering. Friedkin and Johnsen (2011) conducted experiments in which they manipulated network topology for small groups and measured initial opinions, thereby enabling the quantitative comparison of experimental and model results. Although the communication topology was controlled, the network weights assessing interpersonal and self-influence in the model had to be calculated for each group separately on the basis of subjects' post-discussion ratings of interpersonal influence, thereby limiting predictive capability. However, their work remains the most extensive experimental investigation of an opinion network model. More recent work has employed agent-based modeling to qualitatively support and extend experimental results (Mäs & Flache, 2013; Moussaïd, Brighton, & Gaissmaier, 2015; Moussaïd, Kämmer, Analytis, & Neth, 2013).

INTEGRATED EMPIRICAL AND COMPUTATIONAL INVESTIGATION OF GROUP POLARIZATION

This section provides an illustration of how experiment and opinion network modeling can be integrated as applied to group polarization and serves as a prelude to the description of the integrated approach in the next section. Recent research is described that demonstrates how a modeling-oriented approach can synthesize previously disjoint phenomena, generating a novel theoretical explanation of a classic social influence phenomenon, which furthermore predicts an effect unanticipated by existing theory (Gabbay, Kelly, Reedy, & Gastil, 2017). The basic theory is implemented in a simple aggregation model that integrates group polarization with the fundamental social influence process of majority influence and conformity. Further, a model of opinion network dynamics shows how this basic process can arise from a lower-level attitude change framework, that considers how persuasive messages shift both opinion and its associated uncertainty. Both models not only qualitatively agree with the results of an online discussion experiment but, in accord with the proposed integration of modeling and experiment, are in quantitative agreement with the data as well.

Group Polarization Theory

The question of how groups shift toward more extreme positions has been a focus of both traditional social influence research and opinion network modeling although the explanatory mechanisms favored by each are disconnected. Group polarization is said to occur when, in a group composed of individuals already on the same side of an issue, the post-discussion mean opinion shifts further in support of that side as compared with the pre-discussion mean (Brown, 1986; Isenberg, 1986; Myers, 1982; Sunstein, 2002)} . Note that, contrary to common parlance, "polarization" here refers to movement toward one pole rather than divergence toward opposite poles. The seminal experiments in the 1960s focused on ``choice dilemmas" in which subjects were presented with hypothetical scenarios involving the choice between a risky but higher payoff option over a safer, lower payoff one (Brown, 1986). Subjects were asked to choose the minimum odds of success they would accept in order to pursue the riskier option. For most choice dilemma items, discussion led groups to choose lower odds of success as measured by the difference in the group pre and post-discussion means. The effect, therefore, was originally coined the “risky shift.” However, some choice dilemma items tended to produce shifts toward greater caution while others produced no shift in either direction. Cautious items were marked by a very large stake such as someone's health or marriage whereas risky items tended to offer a large potential gain for a small stake. Experiments on group betting involving real rather than hypothetical stakes also have shown a mix of risky and cautious shifts (Isenberg, 1986).

Beyond the risk context, discussion among similarly-inclined individuals was found to cause more extreme social and political attitudes (Keating, Van Boven, & Judd, 2016; Myers & Bishop, 1970; Schkade, Sunstein, & Hastie, 2010). Manipulation of the evidence presented to mock juries exhibited discussion-induced shifts to lower presumed guilt and softer sentences in cases where the evidence was weaker and higher presumed guilt and harsher sentences for stronger evidence (Myers, 1982). Similarly, jury damage awards exhibit polarization (Schkade, Sunstein, & Kahneman, 2000). In general, the contexts in which group polarization occurs are on the judgmental side of the intellective-judgmental spectrum in which purely intellective tasks, such as math problems, have demonstrably correct solutions whereas purely judgmental tasks are matters of personal taste or aesthetics (Laughlin & Ellis, 1986). Most real-world decision contexts such as forecasting and policymaking are characterized by both intellective and judgmental aspects; they may draw on a body of knowledge (e.g., expertise on a country’s political system) yet judgments must be made as to significant uncertainties (e.g., the intentions of political leaders).

Corresponding to the two main pathways of social influence, informational and normative influence underlie the two main explanations of group polarization within social psychology. In the informational route, known as persuasive arguments theory, group members preferentially expose each other to new information in favor of that side. In the normative route, known as social comparison theory, a group norm associated with the broader culture or that particular group’s identity defines a preferred direction on an issue so that opinions shift in the direction of the norm; a norm toward risk-taking, for instance, would lead groups to make riskier choices as a result of discussion. While the informational and normative mechanisms for group polarization have received robust experimental support, they have never been integrated with strong social influence phenomena such as consensus pressure and majority influence. Relatedly, neither explanation has been developed with respect to a clear formal model at the individual group level. Although the informational and normative processes occur at the group level, these theories were operationalized mainly with respect to a population of groups with random initial opinion distributions, over which majority influence could be assumed to cancel out. As a result, group polarization theory is effectively silent as to whether a particular group with a specific initial distribution of opinions will become more extreme. Alternatively, one could make a strong reading of either persuasive arguments theory or social comparison theory that neglects other processes, in which they always predict polarization for homogenous groups (for sufficiently judgmental issues). Either alternative – silence or a uniform prediction of polarization – limits the ability of existing group polarization theory to address real-world contexts such as whether, in the face of a foreign policy crisis, discussion among a country's leadership will induce a shift toward a more extreme course of action.

Opinion network models do not suffer from an inability to go from initial to final opinions since that is their fundamental purpose. The dominant approach to modeling extremism within this literature has been to attribute higher network weights to nodes with more extreme initial opinions (Deffuant, Amblard, Weisbuch, & Faure, 2002; N. E. Friedkin, 2015). This "extremist-tilting" approach is necessitated by the fact that in most continuous opinion models (binary opinions cannot become more extreme), the mean opinion in networks with symmetric weights (i.e., the strength of influence is the same from node *i* to *j* as from node *j* to *i*) remains constant at its initial value therefore preventing the shift in mean exhibited in group polarization. Consequently, extremists must be assigned greater influence over moderates than vice versa in order to shift the mean. Psychologically, this move is attributed to extremists' greater confidence, commitment, or stubbornness. Extremist tilting is not widely accepted within the literature dedicated to group polarization, however, and has received only mixed experimental support (Zaleska, 1982).

Frame-Induced Polarization Theory

This section discusses the frame-induced theory of group polarization introduced by the author, Zane Kelly, Justin Reedy, and John Gastil (Gabbay et al., 2017).[1] This theoretical mechanism is complementary to and can operate simultaneously with the mechanisms of standard polarization theory (shorthand for both persuasive arguments theory and social comparison theory). However, the frame-induced mechanism provides an explanation of group polarization that, unlike standard polarization theory, is integrated with consensus pressure and majority influence, thereby enabling prediction given the group initial opinion distribution. The theory is developed specifically with respect to a quantitative policy under debate, although it should prove applicable to opinions more generally. Examples of quantitative policies include budgets, investment amounts, interest rates, jury damage awards, or military operation sizes. In its emphasis on how the policy is discussed, the theory into account the basis of policy positions and not just

[1] Several terms used in this chapter have been changed from those in (Gabbay et al., 2017): "rhetorical frame" replaces "rhetorical issue"; "distribution reshaping" replaces "rhetorically-induced asymmetry"; and "heuristic frame substitution" replaces "heuristic issue substitution."

the policy value alone. For the particular context discussed here, this basis is grounded in the theory of decision making under risk and uncertainty (Pleskac, Diederich, & Wallsten, 2015), and so constitutes a further theoretical element that is integrated within the frame-induced theory.

Crucial to the frame-induced theory is the recognition of a distinction between the policy under debate and the *rhetorical frame* by which it is discussed. In general, one would expect the rhetorical frame to be a substantive aspect of the policy for which there is substantial disagreement among group members. The frame could represent a key uncertainty or differences in how group members value the outcomes associated with the policy. A given issue may admit multiple frames if there are different dimensions of comparable disagreement. Yet, a single dominant rhetorical frame may emerge due to group-specific dynamics such as deliberate efforts to focus a debate as occurs in political framing (Chong & Druckman, 2007). Persuasion, and hence agreement, is driven by proximity along the rhetorical frame not the policy itself. The shape of the *rhetorical function* that maps policy positions into positions along the rhetorical frame ("rhetorical positions") plays an essential role in frame-induced polarization theory.

Focusing on when uncertainty is the source of the rhetorical frame, uncertainty can generate disagreement if group members have different estimates of the probability of either an unknown variable or an impending outcome. A simple but important example involves a policy that depends on the outcome of a binary gamble so that the likelier one estimates the outcome to be, the more stake one is willing to risk on its occurrence. For instance, one would prefer to invest more in a defense technology company (the policy), the greater one's subjective probability that the pro-defense spending candidate in an election is likely to win (the rhetorical frame). The use of subjective probability of a binary outcome as the rhetorical frame is also relevant to the experiment described below.

Two important behaviors that impact group polarization arise from the distinction drawn between policy and the rhetorical frame: (1) *distribution reshaping*, which preferentially facilitates the formation of extreme majorities and so generates group polarization; and (2) *heuristic frame substitution*, which can enhance polarization on one side of the issue and suppress it on the other. Distribution reshaping arises when a nonlinear rhetorical function causes the relative spacing between group member rhetorical positions to be different than between their policy positions. Consequently, the distribution of group member rhetorical positions will be reshaped with respect to the distribution of policy positions. Such reshaping may reduce the rhetorical distance within some subgroups relative to others as compared with their distances directly along the policy itself, thereby affecting the composition of the majority that emerges during deliberations. Specifically, for a concave (downward curving) rhetorical function, rhetorical position increases more slowly as the policy becomes more extreme. This causes the rhetorical distance between more extreme members to be compressed relative to the distance between more moderate ones. This compression favors the emergence of a majority at the extreme, which then drives consensus to a policy that is more extreme than the mean of the initial policy distribution. For the case where the policy, such as a wager or investment, arises from the subjective probability of an outcome in a binary gamble, Gabbay et al. (2017) show that a concave rhetorical function is expected using the theory of decision making under risk and uncertainty (Pleskac et al., 2015).

As an illustration of distribution reshaping, consider a group of three military planners in wartime tasked with deciding whether to increase or decrease the size of the force allocated to defend a certain territory. The policy is then the change in the number of troops, positive or negative, from the current level. Take a planner's preferred force level to be a function of their estimate of the probability of an enemy offensive against this territory and their assessment of its worth relative to other territories. If there is little disagreement as to worth yet there is fundamental uncertainty as to enemy intentions, then the subjective probability of an enemy offensive is expected to be the dominant source of disagreement and, hence, the rhetorical frame. The rhetorical function is then the transformation that maps a given change in force level to the corresponding subjective probability of an enemy offensive. For instance, say that the three planners are all inclined to boost the force level and that their respective preferences for the increase in troops are (500,1500,2500), which are mapped by the rhetorical function to subjective probability estimates of (0.55,0.65,0.70). These values indicate that the rhetorical function is nonlinear: a policy difference of 1,000 between the first two members corresponds to a change in probability of 0.10 while a 1,000 difference between the second and third members yields a probability change of only 0.05. More specifically, it is concave in that the subjective probability goes up at a slower rate as the increase in force level becomes more extreme. While the policy distribution consists of one member at the mean of 1,500 with the other two an equal distance below and above it, the rhetorical position distribution has one member below its mean of 0.633 and two above; a symmetric policy distribution has been reshaped into an asymmetric rhetorical one. As will be described presently, the theorized opinion change process assumes that the two more extreme members are likely to form a majority, converging at their midpoint of 2,000, which yields the ultimate group consensus.

To yield a systematic tendency for homogenous groups to shift toward the extreme, distribution reshaping must be linked to an opinion change process. The Rhetorically-Proximate Majority (RPM) process put forth in (Gabbay et al., 2017) specifically treats consensus formation. Most group polarization experiments have required consensus and it was the most common outcome in our experiment. Figure 1 illustrates how distribution reshaping combines with the RPM process to generate group polarization. Two separate three-person groups, the *F* and *U* groups (the reason for these designations will be revealed below), are depicted. The policy reference at $x = 0$ is the boundary between the opposing policy sides. If a group is entirely on one side of the policy reference, then it is said to be homogeneous. The *F* and *U* groups are seen to be homogeneous on the positive (pro) and negative (con) policy sides respectively. Both groups have symmetric policy distributions given by $(x_1, x_2, x_3) = (\pm 1, \pm 7, \pm 13)$ where the $+(-)$ sign corresponds to the *F* (*U*) group.

The *F* group is used to describe distribution reshaping and the RPM process. The rhetorical function $\rho(x)$ maps the policy x to the rhetorical position ρ. The rhetorical position scale goes from 0 to 1 and so can correspond to the choice of subjective probability as rhetorical frame. Although the policy of the centrist F_2 is exactly midway between those of the moderate F_1 and the extremist F_3, the *F* group members are arrayed along the shoulder of the rhetorical function so that, with respect to the rhetorical frame, F_2 is closer to F_3 than to F_1; the rhetorical distance $\Delta\rho_{32}$ is much less than $\Delta\rho_{21}$. As agreement is driven by proximity of rhetorical positions, it is therefore more likely that F_2 will join with F_3 to form the rhetorically-proximate majority pair than with F_1. There remains the question of where the RPM pair will converge. Assuming F_2 and F_3 have equal influence on each other, then they should converge at either the midpoint between their rhetorical positions, which is then transformed back to the corresponding policy position, or midway between their policies. Either choice will lead to a policy more extreme than the mean x_2 but the latter process is more direct than the former. In addition, the policy positions are explicitly numerical whereas the rhetorical ones need only be expressed qualitatively. It is therefore more plausible that the RPM pair forms midway between their policies (at $x_{23} = (x_2 + x_3)/2 = 10$). The convergence of the RPM pair on this point is indicated by the solid arrows leading to the open circle. As indicated by the dashed arrow, majority influence then causes F_1, now in the minority, to conform with the F_2, F_3 position. The result is a consensus policy that is more extreme (i.e., further from the policy reference) than the initial policy mean – in accord with definition of group polarization.

Although distribution reshaping due to a concave rhetorical function in combination with the RPM process can explain group polarization, it predicts only a systematic tendency for groups to shift toward the extreme. It is clear from Figure 1 that if F2's initial policy were moved sufficiently close to F1, then those two would form the RPM pair, which would yield a consensus policy below the mean. This ability to predict that individual groups can depolarize against an overall polarization tendency stands in contrast to standard theory's strong prediction of polarization for every homogeneous group.

Note that the references corresponding to the policy and rhetorical frame are offset along the horizontal axis in Figure 1. This misalignment between the policy and rhetorical references can arise from heuristic frame substitution in which a simpler heuristic rhetorical frame is discussed in place of a more complex frame that directly corresponds to the policy. In the binary gamble context, such substitution can occur when there are two distinct gambles that depend on the same random variable but with different thresholds: the policy gamble that directly determines whether one's policy choice is successful and a heuristic gamble that is more intuitively accessible. In a stock investing scenario, for example, the policy gamble could be whether or not the return of the stock over a given period of time would exceed that of a fixed return asset such as a bond. If the stock's return is greater than the bond's return, then the investment is successful. The proper rhetorical frame would then be the subjective probability of that outcome. If it is greater than 0.5, then one should invest in the stock and not the bond. However, discussions about the investment might focus on the more intuitive heuristic gamble of whether the stock's price will rise or fall. Both of these gambles depend on the same random variable, the stock's return, but with different thresholds – the fixed return for the policy gamble, zero for the heuristic gamble – and so they have distinct subjective probabilities. If a probability of 0.5 is taken as the neutral reference for both subjective probabilities, then they will be related to different policy reference points by the rhetorical function.

The *U* group in Figure 1 illustrates the effect of reference point shifting due to heuristic frame substitution. The reference point of the rhetorical function is shifted left from that of the policy itself so that the *U* group members, who are all on the same (negative) side of the policy axis, straddle the rhetorical reference point (U_1 is to the right while U_2 and U_3 are to the left). Because they are arrayed along the roughly linear part of the curve, the *U* group members are subject to weak distribution reshaping. One would expect therefore that the U group would be less prone to polarize than the *F* group. Strict application of the RPM process, however, would, for the case illustrated, lead to the formation of a U_2 and U_3 majority, which would lead to a substantial shift toward the negative extreme as the rhetorical distance $\Delta\rho_{32}$ is slightly less than $\Delta\rho_{21}$. But a small shift in rhetorical positions due to noise or uncertainty could readily flip which distance is smaller, causing the (U_1, U_2) majority pair to form which would lead to depolarization instead. Whether the *U* group will polarize is consequently much harder to predict than for the *F* group. When considered over a population of similar *U* groups, about equal numbers will become more moderate as become more extreme and so there will be no systematic polarization. In contrast, the offset of the rhetorical reference places the *F* group further along the shoulder and so enhances systematic polarization on the positive policy side.

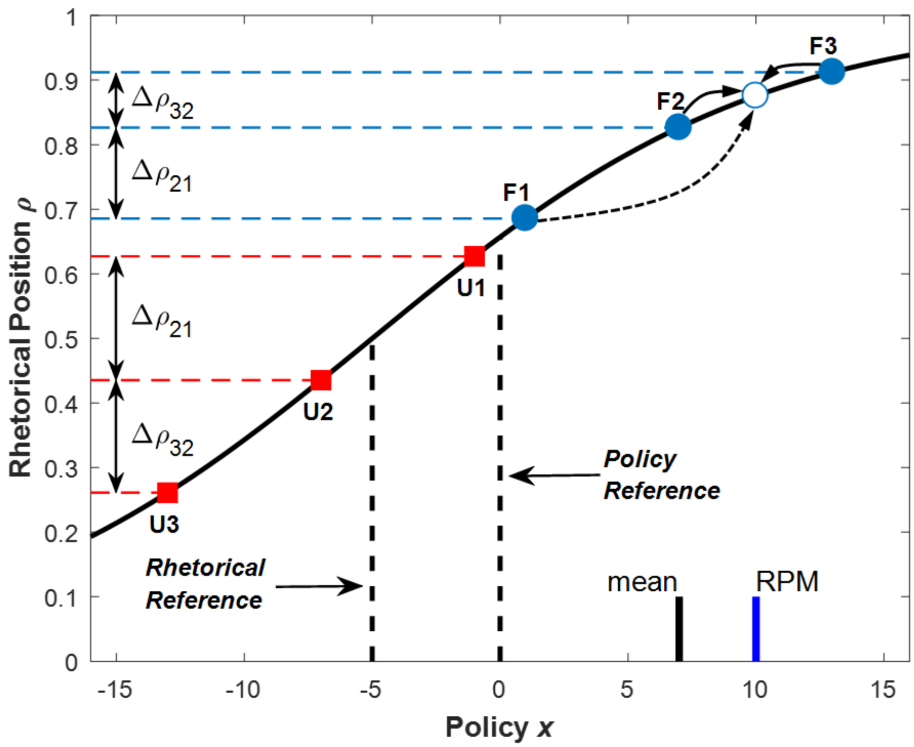

Figure 1. Illustration of distribution reshaping, RPM process, and reference point shifting due to heuristic frame substitution. Rhetorical function relating rhetorical frame position to policy is solid curve. Solid curved arrows indicate formation of the RPM pair by F_2 and F_3; dashed curved arrow indicates conformity of F_1 to form final consensus. Short black line at bottom is the mean of the initial F group policies; short blue line is RPM process prediction for the F group consensus policy.

Accept-Shift-Constrict Model of Opinion Dynamics

Two mathematical models are presented in Gabbay et al. (2017). The RPM process described above is formulated as the RPM model, which determines a consensus policy by a weighted average of the policies of the majority of group members whose rhetorical positions span the least range. Network structure is accommodated by weighting policies by relative node degrees. Rather than static aggregation, the second model describes the opinion change process over time as a result of dyadic-level interactions. This Accept-Shift-Constrict (ASC) model makes two innovations beyond existing continuous opinion network models. First, it makes a distinction between policy (or opinion, more generally) and rhetoric in accord with the theory above. Second, it incorporates a novel uncertainty reduction mechanism which does not require that node uncertainties be visible to others.

The ASC model assumes an underlying dyadic process in which one node sends a message to a receiver node in an effort to persuade the latter. The message can impact both the receiver node's policy and its uncertainty interval around that policy. Conceptually, the model proceeds in distinct "accept,", "shift," and "constrict" phases (although all occur simultaneously in the mathematical formulation). The accept and shift phases occur in the equation that governs the rate of change of the node's policy. In the accept phase, the ASC model assumes that the probability that the receiver node will accept the message as persuasive decreases as a Gaussian function of the *rhetorical* distance between the sender and receiver nodes. The uncertainty of the receiver's position is taken to be the standard deviation parameter in the Gaussian. If a message is accepted, then, in the shift phase, the receiver shifts its policy in the direction of the sender's by an amount proportional to their *policy* difference. The constrict phase is governed by a second equation for the rate of change of a node's uncertainty, modeling a process in which interaction with others with close positions reduces uncertainty. If the sender's rhetorical position is within the uncertainty interval of the receiver, then the receiver decreases its uncertainty but not below a certain minimum value. Accordingly, unlike other models that involve uncertainty dynamics (Deffuant et al., 2002), it is the difference in (rhetorical) positions among dyad members rather than their difference in uncertainties that drives uncertainty change. The network weights in the ASC model represent the influence of one node upon another due to factors such as communication rate and expertise; they need not be symmetric. The ASC model is implemented in terms of coupled nonlinear ordinary differential equations, with two equations for each group member, one for the policy and one for the uncertainty.

A sample simulation of the ASC model as applied to a triad is shown in Figure 2a. All the nodes are connected and the network weights between nodes are symmetric and all equal. The initial policies are set such that x_2 is closer to x_3 than x_1. We observe that x_2 and x_3 form a majority and their uncertainties (λ_2, λ_3) quickly reach their minimum values while that of the minority member (λ_1) stays at its initial value. Consequently, x_1 is more open to accepting messages from the majority pair than vice versa, so that x_1 essentially comes up to the majority position, resulting in a consensus policy that is shifted upward from the mean. This ability for a majority to emerge and persist in the face of minority influence is a crucial part of the frame-induced mechanism of group polarization. This dynamic is not present in linear opinion network models such as the DeGroot, Friedkin-Johnsen, and consensus protocol models. As shown in Figure 2b for the consensus protocol (also known as the Abelson model, a continuous time equivalent of the DeGroot model), no interim majority emerges as all nodes converge simultaneously on the initial mean.

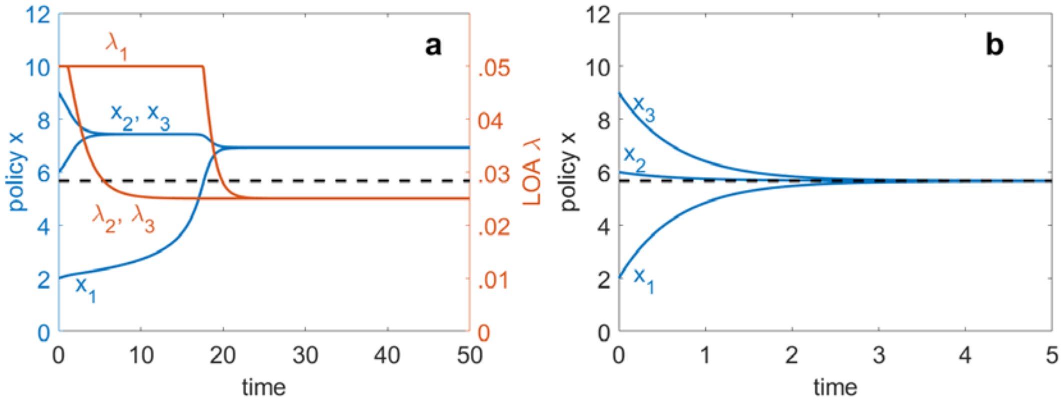

Figure 2. Evolution of policy positions and uncertainties for completely-connected triad. (a) ASC model. (b) Consensus protocol. Dashed lines indicate initial mean policy.

Experiment and Results

In the experiment reported in Gabbay et al. (2017), triads of knowledgeable fans of the National Football League (NFL) wagered on the outcomes of upcoming NFL games. In accordance with standard NFL betting practice, the wager did not concern simply which team would win the game but rather the margin of victory. Professional oddsmakers set an expected margin of victory, known as the point spread, by which the favorite team (the team deemed likely to win the game) is expected to win over the underdog. A bet on the favorite is successful if the favorite wins by more than the spread; otherwise, a bet on the underdog is successful (neglecting actual practice of returning bets when the favorite's victory margin equals the spread). The spread is set so as to estimate the median margin of victory if the game were repeated many times and, empirically, the chances of favorites and underdogs with respect to the spread are effectively equal. As a consequence, the payoff is the same regardless of which team one bets on.

Several days before the selected game, subjects drawn from an online labor pool were asked in a survey to choose which team they expected to win against the spread and how much they would wager from \$0 to \$7 on their choice. Triads were then assembled into online discussion groups according to three manipulated variables: (1) Policy side conditions of "favorite" or "underdog" corresponding to the team that all group members had chosen (there were no groups of mixed team choice). (2) Disagreement level conditions of "high" (\$7) and "low" (\$3,4,5) corresponding to the difference between the highest and lowest wagers in the group. (3) Network structure conditions of "complete" in which all group members could send messages to each other and "chain" in which the intermediate wager person was the middle node and the low and high wager individuals were the ends. Groups discussed the game for 20-30 minutes after which members selected their final individual wagers. The winnings from successful wagers were donated to a specified charity.

Analyzing only groups which made consensus wagers (the vast majority of outcomes), the polarization metric is the difference between the mean post and pre-discussion wagers of each group. If the post-discussion mean is higher, the group is said to have displayed a risky shift. Statistically significant results were found for all three manipulated variables: (1) favorite groups displayed a large risky shift whereas underdog groups showed a small shift not statistically distinguishable from zero; within the favorite groups, a greater risky shift was observed for (2) high disagreement groups than low and (3) complete networks than chains.

These results are not readily explained by either standard polarization theory or existing opinion network models. The differential polarization behavior in particular stands at odds with the informational and normative explanations, which predict that a risky shift should occur for both the favorite and underdog conditions. Since both favorite and underdog groups were homogeneous with respect to policy side, group members have more novel arguments in support of their team choice, and so persuasive arguments theory predicts polarization for both favorite and underdog groups. Similarly, given the low stakes of the task, social comparison theory predicts that a norm favoring risk-taking should be present in both groups and so both should exhibit a risky shift. This differential polarization result is also counter to the extremist-tilting explanation of opinion network modeling because, presumably, a high bet on the underdog is equally as extreme as the same amount bet on the favorite and so the level of extremist tilting is the same for both sides. With respect to the results for disagreement and network structure, their effects have been under-theorized and under-explored in the literature. The only previous experimental investigation of network structure in group polarization found no effect of topology (Noah E. Friedkin, 1999).

The experimental results, however, are in qualitative agreement with frame-induced theory. The differential polarization by policy side arises from heuristic frame substitution in which the question directly related to the wager policy – who will beat the spread? – is replaced by the heuristic one – who will win the game? The subjective probability of the favorite winning the spread is the correct rhetorical frame but is replaced by the subjective probability of the favorite winning the game. While professional gamblers may be able to think directly in terms of the spread victor probability, the game victor probability is a much more natural one for most knowledgeable fans to consider and so constitutes the rhetorical frame operative in the discussion. This substitution also entails a shift of reference point since both gambles depend on the same random variable, the margin of victory, but with different thresholds for their resolution. The reference margin of victory for the spread victor gamble is the point spread whereas the game winner gamble has a reference of zero points. These different references for the margin of victory yield different policy and rhetorical reference wagers respectively. The rhetorical reference is obtained by considering the wager for a subjective probability of the favorite winning the game equal to 0.5. Believing that the game is a toss-up implies that one estimates the margin of victory to be zero and so one should bet on the underdog if the oddsmakers have set a nonzero spread. Therefore, the rhetorical reference equates to some wager on the underdog. If positive and negative wagers are used to represent favorite and underdog bets respectively, then the rhetorical reference corresponds to a negative wager. Accordingly, the *F* and *U* groups in Figure 1 can now be seen as analogous to the favorite and underdog groups in the experiment.

That favorite groups with high disagreement are expected to show a greater risky shift is a consequence of the RPM process described in connection with Figure 1. Expanding the difference between x_1 and x_3 while keeping x_2 fixed implies that the F_2 and F_3 will form the RPM pair at a more extreme policy since x_3 is more extreme. The greater polarization for complete networks vs. chains is due to the greater relative communication rate of the center node in the chain along with its intermediate wager. Rather than forming at their policy midpoint as in the complete network, the greater influence of the chain center node causes the RPM pair to form at a policy that is closer to x_2 and so implies a lesser shift to the extreme.

As a visual comparison between the data and models, Figure 3 displays the observed and simulated pre-to-post discussion shift in the group mean wager as a function of the wager difference (averaged over all groups at each difference) where favorite and underdog groups are shown, respectively, on the positive and negative sides of the horizontal axis. Groups are simulated using their actual wagers and spreads. The weights in the complete and chain networks are set by a priori considerations of the topological effects upon communication rates. In the complete network, all weights are equal whereas the middle node in the chain has twice the weight of the end nodes (these expectations are in approximate agreement with the measured communication rates). The free parameters in the models, the level of risk aversion plus the initial and minimum uncertainties (for ASC), were chosen so as to minimize the total χ^2 error over both networks between the observed and simulated data. The data displays the observed greater polarization for favorite groups, high disagreement level, and complete topology. The RPM and ASC simulations also display these behaviors demonstrating qualitative agreement between the experiment and simulations. That the simulation results mostly pass through the error bars further suggests quantitative agreement whose testing we now discuss.

In general, when statistically testing the fit of a model, one assesses whether its predictions are consistent with the data in the sense that it is reasonably probable that the model could have produced the data given the presence of unmodeled noise. The null hypothesis is that the model is correct, and so support for the model is found if the null hypothesis cannot be rejected. Note that this criterion is opposite to that used in the testing of qualitative effects where one seeks to reject the null hypothesis in order to claim support for the theorized relationship. If the model passes the test, one can only claim that it is consistent with the data not that it is the true model whereas failure to pass the test indicates that the model can be rejected as false. Free parameters – those that cannot be determined without using the dependent variable –

are fit so that the error statistic is minimized. Using more free parameters, however, has the effect of decreasing the maximum allowable error beneath which one can claim support for the model. This makes it harder for models with more parameters to pass the test. However, if two models – a parameter-lean one and a parameter-rich one – both pass the test, one cannot claim more support for one than the other on the basis of the test itself. While the parameter-lean model might be preferred by virtue of its parsimony, the parameter-rich model may be preferred if it has more general scope beyond the experiment under study or if there are more grounds for its causal process in the relevant literature.

The RPM and ASC models were evaluated using a χ^2 goodness-of-fit test. A χ^2 goodness-of-fit test uses the sum of the squared errors between the observed and predicted data points (normalized by the variance at each point) as its test statistic. A threshold of $Q \leq 0.2$ was chosen for rejecting the null hypothesis that the model is correct. This threshold is conservative with respect to the standard significance threshold (p-value) of 0.05 used in testing qualitative hypotheses in that a higher Q-value makes it more difficult for the model to pass the test. The RPM model (one free parameter) has $Q = 0.61$ and the ASC model (three free parameters) yields $Q = 0.3$. Accordingly, both models were found to be consistent with the data. On the other hand, several alternative models, such as the median, a proximate majority model based directly on the policy (not the rhetorical frame), and the consensus protocol, did not pass. Consequently, this test provides a statistical basis for rejecting these models as explanations of the experimental data.

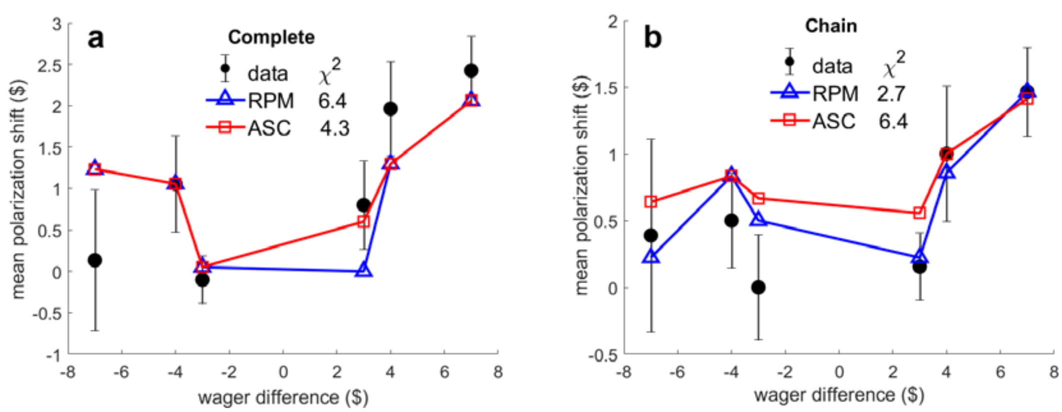

Figure 3. Comparison of experimental data and simulations of ASC and RPM models. (a) Complete network. (b) Chain. Positive and negative sides of the x-axis correspond to favorite and underdog groups respectively. Error bars are standard errors. Simulation results are rounded upward to nearest dollar.

While the agreement between the experiment and the models is encouraging, more experimentation is needed to judge the validity of the theory. Of greater relevance for present purposes is that this work illustrates some of the themes of the integrated approach. First, it demonstrates that quantitative agreement between computational models and experiment is possible. Plots such as Figure 3 comparing data points with error bars against model predictions as a function of an independent variable are common in the physical sciences, but not so in social influence research or social science more generally. The work also shows that the modeling goal of predicting outcomes for groups with specific initial conditions drives a heightened concern for the synthesis of concurrent effects, as done here for group polarization and majority influence. This concern for synthesis can lead to new theory, which, in turn, can predict fundamentally novel phenomena such as differential polarization due to frame substitution. In addition, the synthesis driven by modeling can yield clear predictions for the effects of variables left ambiguous in qualitative theory, such as disagreement level and network structure.

INTEGRATED APPROACH

The proliferation of opinion network modeling research that is largely uncoupled from empirical support may be impeding rather than advancing computational social science. Particularly in fields outside of social science such as physics, computer science, and engineering, an initial model can be subjected to ever more sophisticated analysis or variation, thereby achieving prominence incommensurate with its level of empirical support. This proliferation makes it hard for application-oriented researchers to choose among various models as model popularity need not indicate empirical validity. Likewise, it is difficult for prospective operational consumers of simulations that incorporate opinion network models to evaluate their validity.

For its part, social psychology is not wanting of experimental investigation of social influence. Commonly, a new behavior is discovered experimentally and further experiments are aimed at generalizing and elaborating upon the origins and circumstances of its occurrence. Experiments can extend the behavior to new contexts and winnow down competing theoretical explanations, but much less effort is dedicated to synthesizing it with other social influence phenomena. This lack of synthesis hampers application to real-world contexts. Many experiments have been conducted on group polarization since the 1960s, but the results have not been firmly integrated with majority influence, consensus pressure, or attitude change research bearing on disagreement. As applied to a natural group decision making context for an issue with a substantial judgmental element, the normative influence theory, for example, would advise identifying whether a culturally-salient norm is present that would push groups further to the extreme. But then, given such a norm, it would always predict polarization. Another example is how the hidden profiles paradigm (Schulz-Hardt & Mojzisch, 2012) has not been reconciled with the persuasive arguments theory of group polarization: the former emphasizes how group discussion centers on pieces of information held in common to the detriment of the sharing of unique information held by individuals while the latter hinges upon group members exchanging their unique information.

An approach that centers on integrating experiment with the quantitative testing of computational models of social influence will place models on more solid empirical footing. Of particular importance to this approach are complex systems models (e.g, the ASC model), in which the variables of interest, such as group member opinions, evolve from initial conditions as a result of endogenous feedback with each other and perhaps exogenous signals (Gabbay, 2014). Unlike standard statistical models, complex systems models do not directly posit some generic functional form, often linear or quadratic, that expresses an overall relationship between independent and dependent variables. Rather, the overall functions relating input and output are determined by the unfolding of the processes in the model and so need not result in a function that is expressed in a simple analytic form. However, the predicted relationship is more specific than simply saying, for instance, that there is an interaction effect between variables as in regressions. That complex systems models must deal with specific group initial conditions, rather than an overall population of groups, provides a greater motivation to synthesize different effects than do statistical approaches.

Group polarization research provides an example of how seeking to model specific initial conditions can drive synthesis. One can statistically analyze the overall group polarization in a population of homogeneously-inclined groups with random initial preference distributions without worrying about majority influence. Since groups with a majority preference above the mean would be roughly balanced by those whose majority was below the mean, (preference-based) majority influence could not be the cause of any observed net shift toward the extreme. However, majority influence strongly affects whether a particular group will polarize or not as it can operate counter to polarization when a group has a majority below the mean. This implies that majority influence cannot be ignored when one seeks to predict whether individual groups will polarize. Moreover, the focus on making predictions for specific initial preference distributions, rather than a population of random distributions, helped spur the frame-induced theory's reconceptualization of majority influence as operating over the rhetorical frame, thereby becoming integral to the group polarization process.

Implementation of an integrated approach will entail changes for both the experimental and modeling sides. For the former, the major change is that experiments be designed from the outset with the objective of quantitatively testing computational models not just qualitative hypotheses. For hypothesis testing, one often coarse grains the values of an independent variable in order to construct experimental cells corresponding to a binarization of that variable, as we did with disagreement level above. Quantitative testing of computational models, however, typically requires greater resolution and experiments should be designed so that a sufficient number of variable values are present to conduct a goodness-of-fit test. As complex systems models can utilize specific experimental initial conditions for individuals or groups instead of treating them as random error, greater attention should be given to the distribution of initial conditions in the design than is needed when the goal is simply to populate binarized condition cells.

Beyond a focus on initial and final states, the testing of complex systems models would also benefit from experimental measurements over time, when such measurement is feasible and does not unduly interfere with the process under study.

For the modeling side, a greater focus on developing models capable of being estimated from experimental data is needed. Making quantitative contact with experimental results requires more disciplined consideration of parameters than when more simply endeavoring to show a qualitative correspondence with the data. The temptation to develop rich models must be tempered against the need to estimate parameters from the data itself if they cannot be independently determined. Discretion should be exercised when considering the addition of parameters beyond those directly related to the variable of primary concern. The essential parameters in continuous opinion network models will involve distance in the opinion space.

Another important element is that models be capable of prediction. This requires that, once parameters are estimated on an in-sample population, models are then capable of predicting cases held out of the sample or from a new experiment. As noted above, the network weights in the Friedkin-Johnsen model have primarily been calculated using post-discussion ratings of influence by the group members themselves. Such a procedure precludes prediction. However, it would be possible to test the extremist-tilting explanation of group polarization in conjunction with the Friedkin-Johnsen model if one were to fit the function relating persuasion resistance to opinion extremity.

The focus on a relatively tight number of parameters that can be determined a priori or estimated from the data will make models more robust and applicable across experiments falling into the same broad context. A more ambitious goal of an experimentally-oriented modeling program would be to bridge different experimental contexts, such as problem solving, forecasting, policymaking, and ideological attitudes. As an example, the distinction between intellective and judgmental tasks is an important one in group decision making. At opposite ends of the intellective-judgmental spectrum, purely intellective tasks like math puzzles have solutions that are demonstrably correct whereas purely judgmental tasks are matters of personal taste. Forecasting problems, for instance, lie in between, having intellective elements that are demonstrably right or wrong, such as the record of a football team or what party has the most registered voters in a given district, but also judgmental ones involving the factors likely to be most important in a particular circumstance, such as motivational differences between teams or the impact of national-level political considerations on a local election (Kerr & Tindale, 2011). Where a task lies on the intellective-judgmental spectrum affects the relative importance of social influence effects such as minority or majority influence. Rather than simply categorizing a context as intellective or judgmental and choosing a model accordingly, it would be preferable to define a parameter that gauges the balance between intellective and judgmental factors and thereby the weight of the dynamical mechanisms at play in a given context.

Experiments could then test whether models integrated using the parameter could successfully make predictions for various tasks along the intellective-judgmental spectrum. As an illustration, one might conjecture that with respect to group polarization, the frame-induced mechanism might best suit forecasting problems whereas political ideologies, falling further on the judgmental side, might best be modeled by extremist tilting. Policy making might fall in between the two and a parameter reflecting that balance could then be used to weight relative strengths of the frame-induced and extremist-tilting mechanisms.

A potential hazard of orienting experiments toward model testing is that experiments may be treated as primarily data-fitting exercises in which researchers test a raft of different models driven more by the various mathematical or simulation possibilities rather than by theory. This tendency will lead to models that are narrow in scope and not readily generalized to new circumstances. The SDS literature suffered from this tendency, as mentioned above, and never produced a compelling account of group polarization. However, the increasing prevalence of requirements among journals to make datasets available may help counter the lack of convergence caused by the tendency to seek and emphasize the best fitting models for experiments in isolation. A new model, which best accounts for the results of a particular experiment, can now also be tested against data from previous experiments. The growth of a norm toward testing models against new and old data will encourage the development of more general models.

Initial successes in model development and testing will eventually lead to the emergence of a self-sustaining research community dedicated to the integration of modeling and experiment (an example of a new and virtuous epistemic culture as described in the chapter by Davis and O'Mahony ???). In the short term, a comprehensive program aimed at providing experimental data to test and develop a range of models would help generate the nucleus of such a community as well as advancing social influence research itself. The goal of the program would be to develop general models that integrate different social influence phenomena over a range of contexts rather than the current practice that investigates behaviors in divergent research streams. To effect such synthesis, the program could unfold in phases in which experiments and models initially focus on relatively narrow phenomena with later phases becoming successively more integrative. This would encourage the development of more general and robust models and counter the tendency toward one-off model fitting. Follow-on research could test the models on real-world contexts of interest. Ideally, experiments would be conducted by separate teams of researchers who would then share the data with modeling teams. However, as there is little tradition in social psychology (and social science more broadly) of publishing experimental results without at least some theoretical embellishment, experimental teams could be allowed to develop their own models or included as co-authors on initial publications using their data.

While the goal of quantitatively testing computational models of social influence is challenging, the rewards for doing so would be high. On a scientific level, it would make the study of social influence more synthetic and cumulative. The bar would be raised for evaluating competing theories: a theory implemented in terms of a model that provided a quantitative account of experimental results would be preferable to one that only provided a qualitative account. In addition, the present practice of testing hypotheses experimentally provides guidance on behavioral expectations based on the coarse graining of variables. This practice results in nominal categorizations that can be difficult to extend to more general conditions and synthesize when competing effects and nonlinear interactions are present. Standard group polarization theory, for example, only applies when all group members have common inclinations. Although a more relaxed condition of “most” group members is often stated, that is too vague to enable prediction in situations where groups have members with opposing preferences. Effectively, the reference point is a wall beyond which standard polarization theory is silent. By allowing for gradations of effects, computational models are less constrained by such ambiguous categorizations; for instance, the rhetorical reference point in the ASC model is simply a parameter that affects the rhetorical distance between group members, a distance that can be calculated regardless of whether it spans the reference point or not. As a result, the ASC model can treat the combined *F* and *U* groups in Figure 1 standard theory cannot. This freedom from dependence upon categorizations implies that models can be more readily extended to variable and parameter regimes not yet explored experimentally. In combination with the ability to probe the effects of nonlinear interactions via mathematical analysis and simulation, models can therefore be used to reveal novel, potentially counterintuitive behaviors not anticipated by qualitative theorizing.

Greater incentive to perform replication experiments would be another scientific benefit if experiments were to become more oriented toward computational models. Historically, there has been little incentive for social scientists to perform replications of previously reported effects and for journals to publish them. In the physical sciences, however, better measurements of model parameters, such as physical constants, are valued even if no new effect is reported, as such measurements improve the accuracy and precision of model predictions. Similarly, social influence experiments aimed at testing models would yield improved parameter estimation, and so hold more value than merely replicating an effect. New experiments could repeat earlier ones, but with higher resolution or an extended variable range thereby enhancing the precision and robustness of parameter estimates. It is also possible that systematic deviations from model predictions could be observed pointing the way toward new theory and model development. A greater ability to publish such discrepant experimental data as valuable in its own right (without theoretical explanation) would thereby allow social scientists to learn more from data than is presently the case. Fundamental advances in physics have occurred because of the publication of experimental findings that ran counter to accepted theoretical models. A pivotal event in the genesis of quantum mechanics, for example, was the discovery of the photoelectric effect, a phenomenon at odds with classical physics, which Einstein eventually explained. The concentration on developing general models that minimize the number of free parameters will also discourage data dredging practices in which researchers sift through a large number of covariates in order to find statistically significant, albeit likely spurious, relationships.

The approach outlined above is intended to develop a community within social influence research in which theory, modeling, and experiment proceed in a fashion similar to the physical sciences, albeit with reduced expectations of predictive power. It is not a call to end the traditional paradigm for the investigation of social influence. Although we have argued that the integrated approach will lead to new theories and discoveries, taking it further by demanding that novel theories be implemented formally before experimental testing would, on net, likely hamper the discovery of new behaviors, given the richness of social systems. A more desirable outcome would be for the model-oriented and traditional approaches to work in tandem. The standard testing of qualitative hypotheses could explore variables and effects not yet incorporated within quantitative models. This exploratory role would identify promising areas that could benefit from modeling and facilitate model development by narrowing the range of viable theoretical explanations. It is also possible that some behaviors will not be amenable to quantitative modeling and so remain in the province of qualitative theory in which modeling continues to play its more usual historical role in support of hypothesis generation.

CONCLUSION

Much of the recent surge of activity in computational social science has revolved around the analysis of massive amounts of data available from naturally-occurring activity on the internet and social media involving large networks consisting of thousands or millions of individuals. However, such studies do not shed light on the small group context, which is central to decision making in leadership groups as well as the political attitude change among ordinary citizens. Accordingly, the agenda put forth here emphasizes experiments with human subjects. Given their ability to control conditions, experimental studies can more directly test opinion models than can data from online networks or other observational sources. Network topology and initial opinion distributions can be controlled, the latter enabling testing of the core objective of modeling how opinions change from their initial values, rather than predicting final distributions on the basis of assumed initial conditions. Moreover, experimental results can provide a sounder basis for application of opinion network modeling to large systems. Such applications typically use models based on dyadic or other local interactions and if those models cannot predict the results of small group experiments or be derived from approximations of models that can, then little rationale exists for their use on large systems.

The goal of quantitatively testing computational models of social influence is no doubt an ambitious one. The approach advocated in this chapter centers upon the conduct of experiments explicitly designed to test the quantitative predictions of models rather than the standard experimental paradigm of testing qualitative hypotheses. Its aim is the development of models that can account for a range of phenomena and experimental results. Elements of this approach include: exercising discipline and discrimination with respect to model parameters; conducting goodness-of-fit tests; more highly resolved initial variable conditions; more deliberate control of initial opinion distributions; measuring opinions or other variables over time; greater use of out-of-sample prediction; testing models on new and old data to foster model convergence not proliferation; and parameterizing the nature of group tasks along a spectrum rather than ambiguously assigning them to nominal categories such as intellective or judgmental.

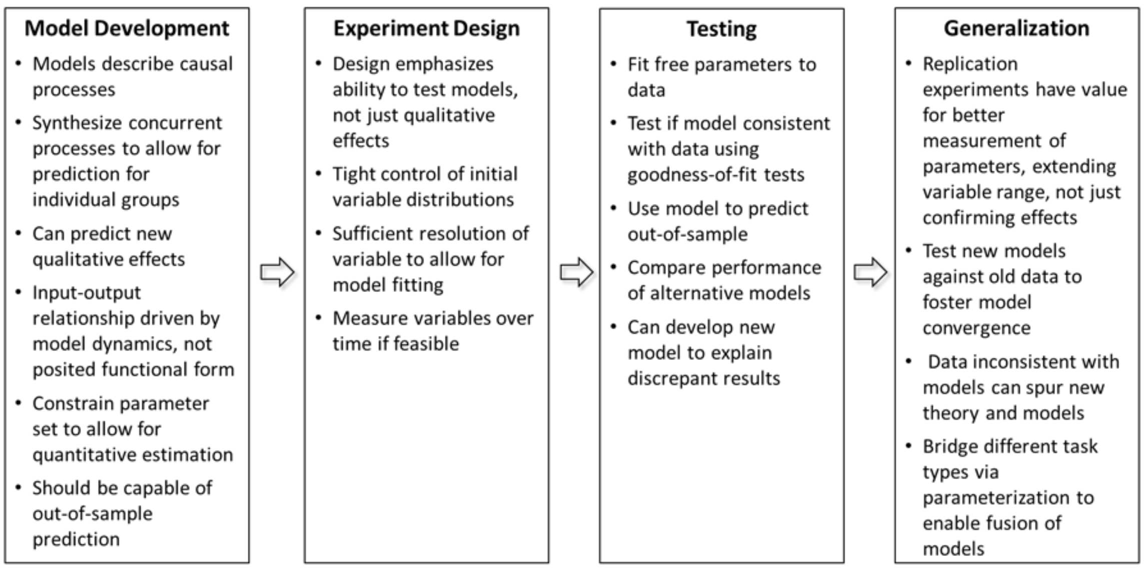

Figure 4. Overview of integrated modeling-experiment approach.

While experimentation with human subjects is much more expensive and laborious than modeling and simulation, a greater emphasis on their integration will enhance both the influence of computational social science and the science of social influence. A major advantage of this integrated approach is an improved ability to synthesize different effects. Since opinion network models make predictions for specific groups, they must take more serious account of effects concurrent to the one under study, which otherwise might be assumed to wash out in a population of groups. Models can synthesize multiple effects more readily than combining different, often ambiguous, categorizations of conditions. The bar will be raised for the evaluation of rival theories with higher precedence given to theories whose associated models are in quantitative accord with experiment. Stronger incentive to conduct experiments for the purpose of providing better measurement or expanding the range of model variables –not just to test hypothesized relationships or competing theories – will be fostered under this approach. Greater replicability will ensue as will the ability to publish anomalous findings thereby spurring new theory and model development.

Ultimately, on an applications level, the integration of quantitative model testing and experiment will raise the confidence and scope with which models can be applied to natural situations for purposes of both prediction and designing interventions to shape outcomes.

Acknowledgments

This research was supported by the Office of Naval Research under grants N00014-15-1-2549 and N00014-16-1-2919.

REFERENCES

Becker, Joshua, Brackbill, Devon, & Centola, Damon. (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences, 114*(26), E5070-E5076. doi: 10.1073/pnas.1615978114

Brown, Roger. (1986). *Social Psychology* (2nd ed.). New York: Free Press.

Burghardt, Keith, Rand, William, & Girvan, Michelle. (2016). Competing opinions and stubborness: Connecting models to data. *Physical Review E, 93*(3), 032305.

Castellano, Claudio, Fortunato, Santo, & Loreto, Vittorio. (2009). Statistical physics of social dynamics. *Reviews of Modern Physics, 81*(2), 591-646.

Chong, Dennis, & Druckman, James N. (2007). Framing Theory. *Annual Review of Political Science, 10*(1), 103-126. doi: doi:10.1146/annurev.polisci.10.072805.103054

Crandall, David, Cosley, Dan, Huttenlocher, Daniel, Kleinberg, Jon, & Suri, Siddharth. (2008). *Feedback effects between similarity and social influence in online communities*. Paper presented at the Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, Las Vegas, Nevada, USA.

Deffuant, Guillaume, Amblard, Frederic, Weisbuch, Gerard, & Faure, Thierry. (2002). How can extremism prevail? A study based on the relative agreement interaction model *Journal of Artificial Societies and Social Simulation, 5*(4).

DeGroot, Morris H. (1974). Reaching a Consensus. *Journal of the American Statistical Association, 69*(345), 118-121. doi: 10.2307/2285509

Devine, Dennis J. (2012). *Jury Decision Making: The State of the Science*. New York: NYU Press.

Düring, Bertram, Markowich, Peter, Pietschmann, Jan-Frederik, & Wolfram, Marie-Therese. (2009). Boltzmann and Fokker–Planck equations modelling opinion formation in the presence of strong leaders. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science, 465*(2112), 3687-3708. doi: 10.1098/rspa.2009.0239

Eagly, Alice H., & Chaiken, Shelly. (1993). *The Psychology of Attitudes*. Fort Worth, TX: Harcourt College Publishers.

Fortunato, Santo, & Castellano, Claudio. (2007). Scaling and Universality in Proportional Elections. *Physical Review Letters, 99*(13), 138701.

Friedkin, N. E. (2015). The Problem of Social Control and Coordination of Complex Systems in Sociology: A Look at the Community Cleavage Problem. *Control Systems, IEEE, 35*(3), 40-51. doi: 10.1109/MCS.2015.2406655

Friedkin, Noah E. (1999). Choice Shift and Group Polarization. *American Sociological Review, 64*(6), 856-875.

Friedkin, Noah E., & Johnsen, Eugene C. (2011). *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*. Cambridge, UK: Cambridge University Press.

Gabbay, Michael. (2007). The effects of nonlinear interactions and network structure in small group opinion dynamics. *Physica A, 378*, 118-126.

Gabbay, Michael. (2014). Data Processing for Applications of Dynamics-Based Models to Forecasting. In J. Egeth, G. Klein & D. Schmorrow (Eds.), *Sociocultural Behavior Sensemaking: State of the Art in Understanding the Operational Environment* (pp. 245-268). McLean, VA: The MITRE Corporation.

Gabbay, Michael, Kelly, Zane, Reedy, Justin, & Gastil, John. (2017). Rhetorically-Induced Group Polarization in Small Opinion Networks. arXiv:1710.10295

Hinsz, Verlin B. (1999). Group Decision Making with Responses of a Quantitative Nature: The Theory of Social Decision Schemes for Quantities. *Organizational Behavior and Human Decision Processes, 80*(1), 28-49. doi: http://dx.doi.org/10.1006/obhd.1999.2853

Isenberg, Daniel J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology, 50*(6), 1141-1151. doi: 10.1037/0022-3514.50.6.1141

Keating, Jessica, Van Boven, Leaf, & Judd, Charles M. (2016). Partisan underestimation of the polarizing influence of group discussion. *Journal of Experimental Social Psychology, 65*, 52-58. doi: http://dx.doi.org/10.1016/j.jesp.2016.03.002

Kerr, Norbert L., & Scott, Tindale. R. (2004). Group Performance and Decision Making. *Annual Review of Psychology 55*(1), 623-655.

Kerr, Norbert L., & Tindale, R. Scott. (2011). Group-based forecasting?: A social psychological analysis. *International Journal of Forecasting, 27*(1), 14-40. doi: http://dx.doi.org/10.1016/j.ijforecast.2010.02.001

Laughlin, Patrick R., & Ellis, Alan L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology, 22*(3), 177-189. doi: https://doi.org/10.1016/0022-1031(86)90022-3

Lorenz, Jan. (2007). Continuous Opinion Dynamics under Bounded Confidence: A Survey. *International Journal of Modern Physics C: Computational Physics & Physical Computation, 18*(12), 1819-1838.

Mäs, Michael, & Flache, Andreas. (2013). Differentiation without Distancing. Explaining Bi-Polarization of Opinions without Negative Influence. *PLOS ONE, 8*(11), e74516. doi: 10.1371/journal.pone.0074516

Moussaïd, Mehdi, Brighton, Henry, & Gaissmaier, Wolfgang. (2015). The amplification of risk in experimental diffusion chains. *Proceedings of the National Academy of Sciences, 112*(18), 5631-5636. doi: 10.1073/pnas.1421883112

Moussaïd, Mehdi, Kämmer, Juliane E., Analytis, Pantelis P., & Neth, Hansjörg. (2013). Social Influence and the Collective Dynamics of Opinion Formation. *PLOS ONE, 8*(11), e78433. doi: 10.1371/journal.pone.0078433

Myers, David G. (1982). Polarizing Effects of Social Interaction. In H. Brandstatter, J. H. Davis & G. Stocker-Kreichauer (Eds.), *Group Decision Making*. London: Academic Press.

Myers, David G., & Bishop, George D. (1970). Discussion Effects on Racial Attitudes. *Science, 169*(3947), 778-779. doi: 10.2307/1729790

Olfati-Saber, R., Fax, J. A., & Murray, R. M. (2007). Consensus and Cooperation in Networked Multi-Agent Systems. *Proceedings of the IEEE, 95*(1), 215-233. doi: 10.1109/JPROC.2006.887293

Pleskac, Timothy J., Diederich, Adele, & Wallsten, Thomas S. (2015). Models of Decision Making under Risk and Uncertainty. In J. R. Busemeyer, Z. Wang, J. T. Townsend & A. Eidels (Eds.), *The Oxford Handbook of Computational and Mathematical Psychology* (pp. 209-231). Oxford: Oxford University Press.

Proskurnikov, Anton V., & Tempo, Roberto. (2017). A tutorial on modeling and analysis of dynamic social networks. Part I. *Annual Reviews in Control, 43*(Supplement C), 65-79. doi: https://doi.org/10.1016/j.arcontrol.2017.03.002

Schkade, David, Sunstein, Cass R., & Hastie, Reid. (2010). WHEN DELIBERATION PRODUCES EXTREMISM. *Critical Review, 22*(2-3), 227-252. doi: 10.1080/08913811.2010.508634

Schkade, David, Sunstein, Cass R., & Kahneman, Daniel. (2000). Deliberating about Dollars: The Severity Shift. *Columbia Law Review, 100*(4), 1139-1175. doi: 10.2307/1123539

Schulz-Hardt, Stefan, & Mojzisch, Andreas. (2012). How to achieve synergy in group decision making: Lessons to be learned from the hidden profile paradigm. *European Review of Social Psychology, 23*(1), 305-343. doi: 10.1080/10463283.2012.744440

Sunstein, Cass R. (2002). The Law of Group Polarization. *Journal of Political Philosophy, 10*(2), 175-195.

Török, János, Iñiguez, Gerardo, Yasseri, Taha, San Miguel, Maxi, Kaski, Kimmo, & Kertész, János. (2013). Opinions, Conflicts, and Consensus: Modeling Social Dynamics in a Collaborative Environment. *Physical Review Letters, 110*(8), 088701.

Zaleska, Maryla. (1982). The Stability of Extreme and Moderate Responses in Different Situations. In H. Brandstatter, J. H. Davis & G. Stocker-Kreichauer (Eds.), *Group Decision Making*. London: Academic Press.

Affiliation: Applied Physics Laboratory, University of Washington, Seattle

Biosketch

Michael Gabbay is a Senior Principal Physicist in the Applied Physics Laboratory at the University of Washington. His research involves the development and application of mathematical models and computational simulations of network dynamics, focusing on social and political systems. Dr. Gabbay's work seeks to advance both basic and policy-relevant research. He has conducted empirical research using political rhetoric, human subjects experiments, and analyst input and has applied his models and methods towards understanding and anticipating the behavior of real-world militant networks and government leadership groups. His publications on political networks have appeared in both academic and policy-oriented venues. Dr. Gabbay has also conducted research in the dynamics of nonlinear oscillators, nonequilibrium pattern formation, and signal processing. He received his Ph.D. in Physics from the University of Chicago and a B.A. in Physics from Cornell University.

Appendix 4

Fratricide in Rebel Movements: A Network Analysis of Syrian Militant Infighting

Fratricide in rebel movements:
A network analysis of Syrian militant infighting

Emily Kalah Gade,
Department of Political Science, University of Washington

Mohammed M Hafez,
Department of National Security Affairs, Naval Postgraduate School

Michael Gabbay,
Applied Physics Laboratory, University of Washington

Abstract

Violent conflict among rebels is a common feature of civil wars and insurgencies. Yet, not all rebel groups are equally prone to such infighting. While previous research has focused on the systemic causes of violent conflict within rebel movements, this article explores the factors that affect the risk of conflict between pairs of rebel groups. We generate hypotheses concerning how differences in power, ideology, and state sponsors between rebel groups impact their propensity to clash and test them using data from the Syrian civil war. The data, drawn from hundreds of infighting claims made by rebel groups on social media, are used to construct a network of conflictual ties among 30 rebel groups. The relationship between the observed network structure and the independent variables is evaluated using network analysis metrics and methods including assortativity, community structure, simulation, and latent space modeling. We find strong evidence that ideologically distant groups have a higher propensity for infighting than ideologically proximate ones. We also find support for power asymmetry, meaning that pairs of groups of disparate size are at greater risk of infighting than pairs of equal strength. No support was found for the proposition that sharing state sponsors mitigates rebels' propensity for infighting. Our results provide an important corrective to prevailing theory, which discounts the role of ideology in militant factional dynamics within fragmented conflicts.

Keywords: Syria, civil war, fragmentation, infighting, social network analysis, ideology

Corresponding author: ekgade@uw.edu

A revolutionary's worst enemy is often another revolutionary.

—Lichbach (1995: 203)

Introduction

Infighting among rebels is a common feature of civil wars and insurgencies. Rebel movements are usually divided into brigades that fight under several factional banners with varying degrees of coordination.[1] This fragmentation generates a competitive landscape in which violent infighting occurs frequently. The history of civil conflicts is replete with dramatic instances of rebel-on-rebel fratricide (Bakke, Cunningham & Seymour, 2012).[2] The ongoing Syrian civil war offers a stark reminder of how rebels can turn on each other while simultaneously waging war against a formidable regime.

The puzzle of rebel infighting can be addressed at either the *systemic* or *dyadic* level. The burgeoning literature on interrebel wars almost exclusively focuses on systemic risks that generate conflictual rebel relationships. These include the problem of credible commitments born out of anarchy (Christia, 2012), the depth of movement fragmentation (Cunningham, Bakke & Seymour, 2012), regime weakness or impending rebel victory (Lichbach, 1995), the presence of lootable resources (Fjelde & Nilsson, 2012), and the quest for patronage within violent patrimonial political systems (Seymour, 2014). In this article, we take a dyadic approach to understanding which groups are most prone to infighting when rebel movements descend into factional conflicts. Investigation at the dyadic level helps us go beyond the systemic assumption of unit homogeneity and thus can offer finer predictions about who is likely to enter into fratricidal wars.

We make three contributions—theoretical, methodological, and empirical. Theoretically, we investigate the effects of *power*, *ideology*, and *state sponsorship* on the propensity for infighting in rebel dyads. We conceptualize power in conflict dyads as either *symmetric* or *asymmetric* (i.e. groups of similar or dissimilar strength, respectively). Whereas power parity may generate conflict between an established rebel faction and a rising competitor, power

---

[1] Of 181 insurgencies since 1946, more than half involved multiple insurgent groups. Since the 1980s, 64% involved multiple rebel factions (Jones, 2017: 168).

[2] Iconic episodes of interrebel fratricide include Stalinists against Trotskyists during the Spanish Civil War (May 1937); Yugoslavian Communist Partisans against the Nationalist Chetniks during World War II; the Algerian National Movement against the National Liberation Front during their war of independence from France (1954-1962); and Al-Qaida against Iraqi Islamists and nationalists during the American occupation of Iraq (2003-2011).

asymmetry may invite rebel aggression by strong factions against their weaker rivals. We test both propositions. We conceive of ideology in conflict dyads as either *proximate* or *distant* (i.e. groups with overlapping ideological positions or opposing ideological preferences, respectively). We hypothesize that higher rates of conflict are likely between ideologically distant groups than between those that are ideologically proximate. Lastly, we explore the potential effects of state sponsorship on interrebel conflicts by looking for the presence or absence of *overlapping state sponsors* in rebel dyads. We posit that rebels that share state sponsors are incentivized by their external patrons to forge unity and will thus experience less infighting than those with distinct state sponsors.

Methodologically, we introduce a network-analytic approach to explain the determinants of salient conflict dyads in rebel movements. Conflict dyads can be embedded within a network of movement infighting, which enables the detection of patterns of infighting relationships. We use assortativity, community structure detection and network simulation, as well as an Additive and Mixed Effects (AME) latent factor model to evaluate the effects of power, ideology, and state sponsorship on generating conflict within dyads. We also run a number of robustness tests including validation of our findings using Exponential Random Graph Models (ERGM), another approach to statistical inference on networks (Desmarais & Cranmer, 2017).

Empirically, we rely on relational and quantitative analysis of the ongoing Syrian civil war, a conflict that is at the center of regional and international insecurity, multiple humanitarian crises, and military interventions by major powers. We constructed a unique database of three years of rebel infighting for the period of January 2013 to December 2015. The data comes from primary insurgent documents such as rebel operational communiques, social media postings, and jihadist web forums. We also drew upon rebel groups' political programs and manifestos to capture their ideological leanings. Lastly, we collected data on power measures, operational location, and state sponsors of the most prominent groups involved in interrebel conflicts. This rich dataset allows us to test our hypotheses with a number of robustness checks to bolster the validity of our conclusions.

We find compelling evidence that ideology is a major driver of infighting in rebel movements. An ideological difference among rebel dyads consistently increases their propensity for infighting, a result that is robust across analyses and time. Specifically, sectarian jihadists were the most prone to engage in inter-rebel wars and they did so mainly with non-sectarian

Islamists and with secular nationalists and Kurdish separatists. Comparatively, groups that were nationalists or Kurdish separatists did not fight among each other as often as groups of different ideological types. We also observe strong, although less consistent, evidence for power asymmetry in fighting dyads; Syrian rebel infighting is usually between groups of disparate strengths. Lastly, we found no relationship in any model between state sponsorship and infighting. Despite the rivalry among the rebel sponsoring states, we could not find clear evidence that this rivalry shaped militant infighting patterns in the Syrian conflict.

Power in conflict dyads

In an anarchic context with no central authority to enforce binding promises within rebel movements, information and credible commitments problems force rebel groups to be self-regarding and consider their survival above all else (Christia, 2012). Relative power considerations can determine who falls victim to aggression, who merely survives, and who ultimately thrives and captures the lion's share of post-conflict spoils (Krause, 2017). Relative power considerations can lead to two predictions about which rebel dyads are more likely to fight. When rebels confront each other, their power distribution can be either *asymmetric* (one group is substantially more powerful than the other) or *symmetric* (both groups are roughly equal in capabilities).[3] Both scenarios are capable of generating interrebel conflicts.

Powerful groups can exploit the asymmetry in forces by eliminating minor players that infringe on their territory and resources (Fjelde & Nilsson, 2012). They may also attack weaker groups that hold the potential to grow in power and thus challenge their leadership in the future (Pischedda, 2018). Strong rebel groups can also target weaker factions that may act as spoilers in conflict-ending negotiations. Although less intuitive, it is possible for weaker actors to undertake the risk of challenging powerful rebel groups because the payoff is quite high if they are successful (Krause, 2017). This 'gamble for resurrection' is especially likely if the minor challenger is going after a powerful group that controls a resource-rich territory, which can rapidly accelerate the ascendency of the weaker party (Fjelde & Nilsson, 2012). Thus, our first power hypothesis:

[3] When referring to power, we mean relative power capabilities as measured by estimates of the group size. Although an imperfect measure, group size is often used in large-N statistical analyses (Akcinaroglu, 2012; Christia, 2012; Krause, 2013/4). We make the assumption that group size is a proxy for other elements of rebel power, such as financial resources.

Hypothesis 1a: Asymmetric power – Infighting will be more likely between groups of disparate power.

In contrast, symmetric power distribution can produce infighting among rebels as an emerging and dissatisfied militant organization approaches parity with an established rebel group. Two equally powerful rebel groups could threaten one another's security and leadership aspirations, so power parity is a cause for concern for established rebel organizations (Krause, 2017). Two mechanisms help explain how power parity can unleash interrebel violence. The disruption to the existing rebel power hierarchy leads to greater conflict as the hegemonic rebel group, feeling threatened by a rapidly rising rebel faction, seeks to prevent the latter's continued ascendancy. Or, the newly ascendant rebel power itself could initiate conflict by challenging the status quo under the hegemonic rebel faction because it seeks greater representation within the rebel institutional hierarchy (McLauchlin & Pearlman, 2012). Thus, our second power hypothesis:

Hypothesis 1b: Symmetric power – Infighting will be more likely between groups of comparable power.

Ideology in conflict dyads

Rebel groups are fragmented along their ideological preferences, not just their power capabilities. Ideology reflects a group's political demands, normative commitments, and future objectives. It also helps bind rebels to their commanders by fostering identification with group goals and it can motivate commitment and sacrifice (Lichbach, 1995: 92-93). That is why insurgent organizations from diverse traditions—Marxists, Maoists, ethnonationalists, and fundamentalists—dedicate time and resources to socialize their recruits ideologically (Oppenheim et al., 2015; Hoover Green, 2016). We would expect that under scope conditions of ideological diversity, competition and conflict will shape interrebel relationships (Seymour, Bakke & Cunningham, 2016).

Following Gutiérrez Sanín and Wood (2014: 215), we define ideology in rebel movements as a:

> systematic set of ideas that includes the identification of a referent group (a class, ethnic, or other social group), an enunciation of the grievances or challenges that the group confronts, the identification of objectives on behalf of that group (political change – or defense against its threat), and a…program of action.

We operationalize this definition along three dimensions: *conflict framing, ideal polity,* and *territorial aspiration*. Conflict framing specifies the primary referent group for which rebels are fighting, and the out-groups they find most threatening. This is particularly important for conflicts with multiple identity groups, which is common in multiethnic civil wars. Ideal polity refers to the nature of the post-conflict political order that rebel groups aim to create. This dimension captures traditional right-left ideological divides as well as divisions between those seeking to create secular or fundamentalists polities. Territorial aspiration refers to the boundaries of the ideal polity, addressing the core debate between those who wish to maintain the territorial integrity of their states and those who seek to break up the polity into multiple states. Movements with shared conceptions of the ideal polity sometimes diverge over the territorial boundaries of that polity.[4] Territorial aspirations have been at the root of many secessionist civil conflicts, resulting in 131 sovereign states coming into existence since 1945 (Griffiths, 2016).

We hypothesize that divergence along these three ideological components can aggravate infighting in rebel dyads. Conversely, group dyads with similar ideological positions along these three dimensions will exhibit low rates of infighting. Thus, our ideological hypothesis:

Hypothesis 2: Ideological distance – The greater the ideological distance between two rebel groups, the higher the likelihood of infighting.

Three causal mechanisms help explain how ideological differences can produce infighting. First, groups with fundamentally divergent conceptions of the ideal polity are likely to view their cohabitation in the rebel field as mutually threatening. Not only do their competing ideological visions violate their core normative commitments for which they are making the ultimate sacrifice, their divergent conflict objectives make credible commitments difficult to

[4] For example, Islamists today are divided between those who favor establishing an Islamic order within the modern national state and those that harbor the irredentist ambition of restoring an Islamic caliphate.

uphold. Absent trust, competing camps see their coexistence as a zero-sum game with little possibility for power sharing in the future.

In contrast, groups with shared conceptions of the ideal polity corroborate each other's core political preferences and thus can readily signal to their ideological kin their intentions to share power in the post-conflict political order. Moreover, the ascendancy of ideologically similar groups is less threatening to one's core constituency and sponsors, reducing the pressure to compete over leadership. Conversely, a conflict between ideologically similar groups can expose the conflicting parties to condemnation from their supporters because their infighting undermines the unity of their ideological faction.

A second ideological mechanism involves the relationship between conflict framing and targeting. Groups with opposing conflict narratives are likely to adopt divergent targeting portfolios (Gutiérrez Sanín & Wood, 2014). For instance, an overtly sectarian conflict frame may justify expansive attacks against the civilians of a rival sect more readily than a frame that rejects sectarian divisions and, instead, paints all of the nation and its diverse sects as equally vested in forging a new polity. The debate over the legitimate targets of violence is often a key source of dissension within rebel movements, inviting open conflict.

A third ideological mechanism relates to competing visions of territorial sovereignty. As a conflict becomes protracted, the territorial integrity of the state may become a subject for negotiation. Rebels that harbor broader or narrower territorial ambitions may clash with rivals that seek to maintain the extant state boundary. Removing spoilers from the rebel movement can thus drive interrebel conflicts.[5]

State sponsorship in conflict dyads

State sponsorship can generate both rebel unity and rebel rivalries. Civil wars invariably invite external actors to intervene on behalf of the combatants, seeking to project influence and prevent rival states from adversely shaping the conditions for conflict termination.[6] The sponsorship of proxy actors is a cost-effective way for states to compete with their state rivals (Salehyan, 2010). External patrons thus provide arms, money, supplies, or sanctuaries to rebel groups in the

[5] For example, the conflict between Hamas and Fatah during the 1990s revolved around the former's refusal to accept a two-state solution. Hamas sought to sabotage the peace process through suicide attacks, which led the Fatah-led Palestinian Authority to crackdown on Hamas (Kydd & Walter, 2002).

[6] According to Jones (2017: 136), of 181 insurgencies between 1946 and 2015, 82% involved outside support.

expectation that these rebels will exhibit sufficient discipline and cohesion to fulfill their patron's strategic aims. Sponsors can threaten to withhold financing and war materiel from those who are jeopardizing a cohesive rebel coalition (Lichbach, 1995: 179). Bapat and Bond (2012) view such external leverage as an important interrebel institution that can help overcome the credible commitments problem, increase cooperation, police against side negotiations, and mediate conflicts between rebel groups.

However, state sponsors can also undermine rebel unity by incentivizing some rebels to challenge their rivals (Tamm, 2016). This is particularly the case when multiple state sponsors with opposing political agendas seek to foster their own proxy clients through patronage. The presence of multiple sponsors increases the degrees of freedom rebel groups can exercise to support themselves and reduces the degrees of freedom any individual external patron can exert to foster cohesive rebel coalitions (Salehyan, Siroky & Wood, 2014). Thus, whereas overlapping state sponsorship in rebel dyads should mitigate conflict, this moderating effect is diminished in dyads with non-overlapping state sponsors.

Hypothesis 3: Overlapping state sponsorship – Rebel groups that derive support from the same state sponsors will experience less infighting than those who have distinct state benefactors.

Network analytic methodology

Social network analysis in political science is designed to account for interdependence within a system of political actors (Ward, Stovel & Sacks, 2011; Hafner-Burton, Kahler & Montgomery, 2009; Maoz, 2010). Dyadic models that assume independence of observations mischaracterize relational data because infighting is dependent upon relationships with a range of groups within the system (Dorussen, Gartzke & Westerwinter, 2016). For example, network analysis can account for how variation among groups in their degree of infighting (their total number of clashes with other groups) constrains the overall pattern of infighting. The relative numbers of high and low degree groups shape the extent to which high degree groups must primarily fight each other or can fight with many lower degree ones instead. The fragmented nature of asymmetric conflicts makes network analysis a promising quantitative approach for evaluating militant behaviors in multiparty wars (Zech & Gabbay, 2016).

A social network consists of nodes and the ties between pairs of nodes. The nodes can be individuals, organizations, or countries and the ties signify relationships such as communication, cooperation, or conflict. In our empirical analysis, we employ a network in which ties represent the number of infighting episodes between group dyads. Four different methods of investigating and testing the relationship between the conflictual network and the independent variables are applied: (1) comparing the 'assortativity' – a measure of the variable's tendency toward producing homophily or heterophily (connection of like or unlike nodes respectively) – of the observed network with the assortativity distribution obtained from a null model simulation; (2) correlating the variable with important patterns in the network as found via eigenvector-based representations of community structure; (3) a simulation of tie formation that explicitly includes the variable to estimate the characteristic zone within which conflict is enhanced (homophily) or suppressed (heterophily); and (4) an additive and multiplicative effects (AME) latent factor model, relying on the Markov Chain Monte Carlo algorithm. The first three are used to analyze power and ideology separately and the fourth allows us to compare power, ideology, and state sponsorship simultaneously.

Assortativity is the standard measure used to assess whether tie formation is driven by similarity with respect to a scalar variable (as we operationalize power and ideology). The assortativity α is the correlation of the variable values at each end of a tie taken over all ties (see Appendix). An α value of +1 corresponds to a network with maximal homophily whereas -1 signifies maximal heterophily. For statistical testing purposes, the assortativity cannot be treated as one would treat a standard correlation because ties are not taken to be independent. Accordingly, we compare the observed assortativity with the distribution obtained from a null model simulation in which the independent variable of interest is not included: if the observed value is greater (less) than the simulation means then the tie formation process exhibits homophily (heterophily).

Network structure can be visualized in a way that relates to the assortativity of the variable of interest. The modularity matrix is a transformation of the tie data (see Appendix) that is often used for community detection purposes (Newman, 2006). Its eigenvector decomposition can be used to identify patterns of tie formation that are shaped by the variable. If tie formation displays homophily with respect to the variable, then the variable should correlate to some extent with one of the highly-ranked (most positive eigenvalues) eigenvectors. However, if tie formation displays heterophily, the variable should instead correlate with one of the lowest-ranked (most negative eigenvalues) eigenvectors (Newman, 2006).

In the null simulation, nodes form ties (fight with each other) probabilistically. Each iteration consists of the placement of a tie between nodes where the iterations proceed up to the total number of ties in the observed network. The simulation seeks to reproduce the observed node degrees and so assumes that the propensity of a group to fight with other groups is known but not the distribution of its infighting ties. As a result, each node can only receive a maximum number of ties equal to its observed degree (it is not always possible to reproduce the degrees exactly, but the differences are typically small). At any given iteration, the degree deficit by which a node's current degree falls short of that maximum affects its tie formation probability – the larger the degree deficit, the more likely it will form a tie. The degree deficit decreases until it reaches zero at which point a node can no longer form ties. For the null model simulation, the probability of tie formation is proportional to the product of their degree deficits only.

To account for homophily in a node variable, the null simulation is modified so that the probability of tie formation also depends on the distance between the node variables. The variable-dependent probability is taken to fall off as a Gaussian function of the distance where a characteristic length scale l defines the preferred zone within which interactions are likely. The heterophily simulation is similar except that now interactions are more likely outside the zone defined by the characteristic length scale, which we refer to as the suppression length l_S. The purpose of these simulations is to see if a simple model of interactions including the variable of interest can minimize the error with the observed network at a well-defined value of the length scale. If so, additional evidence is thus provided for the operation of homophily or heterophily as well as an estimate of the length scale itself. For instance, if a heterophily simulation of ideology yielded a suppression length of $l_S = 2$ then that would imply that a group is more likely to fight with groups that are outside a distance equal to half the range of the full five-point ideological scale we deploy below. The Appendix shows the simulations' mathematical formulation.

We augment these methods with two types of network regression analysis. We use the AMEN package in R to estimate an additive and multiplicative random effects model of militant infighting in Syria (Hoff, 2015). This model allows for the inclusion of both nodal and dyadic covariates. Latent space methods have been applied to international conflict data (Minhas, Hoff

& Ward, 2016). We also use an ERGM method as a robustness check in the *Supplemental Material*.

The Syrian civil war

The Syrian civil war began as a peaceful Arab Spring movement but quickly formed into an armed insurgency against the regime of Bashar al-Assad (Lister, 2015). The conflict further evolved into a sectarian civil war and regional proxy conflict between the Gulf states and Iran (International Crisis Group, 2013). Non-state actors, including foreign fighters and transnational organizations like Al-Qaida and Hezbollah, followed suit. Major powers—Russia and the United States—also intervened to shape conflict outcomes (Phillips, 2016).

Infighting data and variables

We collected data on 508 distinct infighting episodes between rebel groups in Syria from 1 January 2013 to 31 December 2015, yielding 697 INFIGHTING dyadic ties (some episodes involved multiple groups and fronts, each of which was coded as a distinct infighting dyad). Our unit of analysis is the rebel group, which we define as a collection of armed fighters, ranging from several hundred to several thousand men and women, that have a commander and a distinct organizational identity as represented by a logo, and that uses violence in the course of a civil war or an insurgency to achieve publicly stated political aims against an incumbent regime and its allies. When rebel groups fight as part of formal coalitions or joint operation rooms, we disaggregate those broader units into their member groups and distribute infighting ties (dyads) to all the subgroups or to the specific ones involved in the infighting.

Rebel INFIGHTING, our dependent variable, is defined as actual violent interactions between rebel groups. Violent interactions include armed clashes; firing artillery at rival positions; assassinating or executing rivals; arresting rebels or holding them captive; militarily advancing on a rival's territory or checkpoint with the intent of capturing it; and blowing up buildings, headquarters, or checkpoints that belong to one's rivals with car bombs or suicide attackers. INFIGHTING does not include political disputes, defections, expulsions from the group, splintering, or counter-alliances. There were many infighting episodes that went on for days and weeks. Some turned into infighting campaigns that spanned several months. To accommodate such spans of continuous clashes, ties are defined at the month level with a tie between groups assigned for a given month if at least one violent interaction took place (for complete coding rules, see *Supplemental Material)*. In the AME Models described below, we use a square root transformation of infighting counts to approximate a normal distribution, with additional model specifications displayed in the *Supplemental Material*.

We selected 44 rebel groups to track based on think tank and US Government reports regarding the major players in the conflict. Out of these 44 groups, we analyze the 30 that were involved in at least one episode of infighting during the 2013-2015 period. We define the POWER variable to be the medium estimate of a group's number of members. These 30 groups range in POWER from 500 to as many as 40,000 members as shown in Table I in the *Supplemental Material*, which also includes information on ideology, state sponsors, location, and years of existence. While the analysis set of 30 groups is far from exhaustive, we are confident that we have covered the major players involved in infighting.[7]

For Group IDEOLOGY, we hand-coded major ideological statements of the 30 groups that were involved in infighting episodes. We evaluated groups for three ideological areas of relevance to the Syrian conflict. Sectarianism serves as our CONFLICT FRAME variable: groups with high sectarianism scores cast the conflict as Sunnis vs. Shiites (Alawites), whereas groups with low sectarianism scores have little or no anti-Shiite rhetoric. Salafism, which measures the extent to which groups ascribe to that highly puritanical strain of Sunni Islam, provides our IDEAL POLITY variable. The use of Salafism better resolves differences within various stripes of Islamists than a simple secularism vs. Islamism scale. Revisionism is used for the TERRITORIAL ASPIRATION component of ideology: groups with low scores seek to preserve Syria's territorial integrity, whereas a high score signifies a desire to abrogate it, in particular as do Caliphate-minded sectarian jihadists or Kurdish separatists.

A five-point scale was used for each component, and nationalists/Kurdish Separatists tend to fall on the low end (1) and sectarian jihadists on the high end (5) (see Table IV in the *Supplemental Material* for the ideological scores of the 30 groups). In addition, we also constructed an AVERAGE IDEOLOGY variable from the average of CONFLICT FRAME, IDEAL

[7] We compared our groups to the Uppsala Conflict Data Program (UCDP) dataset, which collected infighting episodes at the year level. We believe our data represents an important step forward because it is at the month level and is focused on individual groups rather than alliances of groups. Although here we limit our network to the 30 groups that engaged in infighting, we also substantiated our findings using all 44 groups through both AMEN and ERGM analyses (see *Supplemental Material*).

POLITY, and TERRITORIAL ASPIRATION to serve as an aggregate variable for visualization purposes and to provide an additional test of overall homophily.

Secular nationalists are represented by the Free Syrian Army (FSA), an umbrella organization that has many affiliated brigades. The FSA frames the Syrian rebellion as a national and democratic revolution that encompasses Syria's diverse ethnic and religious communities. It avoids overt sectarianism and rejects the goal of establishing an Islamic state ruled by strict religious laws (International Crisis Group, 2012).

On the other end of the ideological spectrum are groups like Al-Nusrah Front (ANF) and the Islamic State (ISIL). Formed in January 2012, ANF is an outwardly sectarian and jihadist faction that frames the conflict not as a revolution but rather as a religious war against a secular regime ruled by heretical Alawites. It calls for the establishment of an Islamic state governed by strict religious law. In 2013, ISIL broke up the ranks of the ANF to form an even more extreme sectarian jihadist faction. Its goal has been to carve out an Islamic state exclusively for Sunni Muslims that stretches from western Iraq to northeastern Syria.

Residing between the two poles of secular nationalist and sectarian jihadist are many Islamist factions ranging from Muslim Brotherhood affiliates such as Liwa al-Tawhid to Salafists such as Ahrar al-Sham Islamic Movement (ASIM). We categorize these groups as Salafist nationalists because they want to establish an Islamic state within the extant boundaries of Syria's national territory and do not frame the conflict in overtly sectarian terms.

Kurdish communities formed their own combatant organizations, notably the People's Protection Units (Yekîneyên Parastina Gel, YPG), to safeguard their territories from both regime forces and hostile rebels (International Crisis Group, 2014). The secular YPG views Kurdish co-ethnics as its primary constituency, for which it seeks autonomy within, or separation from, the Syrian state.

Figure 1 displays the number of infighting incidents by ideological dyads across time. Infighting increased dramatically across the network in 2014 and 2015. Groups with an ideological difference have more infighting ties in every year and overall than groups with shared ideologies. Jihadist groups also fight among themselves more frequently than do dyads composed of secular nationalists or Kurdish separatists.

[Figure 1 about here]

Rebel infighting appeared in every Syrian governorate, but the vast majority of the infighting took place in the rebel-held areas of Aleppo (38%) and the Damascus countryside (19%), followed by Idlib and Dayr al-Zawr (11% each). It is worth noting that most of Syria's oil and gas resources are concentrated in the eastern part of the country near the Iraqi border—i.e. in Dayr al-Zawr and Hasaka. Rebel infighting in those regions is about 11% and 7%, respectively. Thus, most of the infighting took place outside of the resource-rich regions in the period under consideration.

ISIL fought the most with rival factions; it was involved in 41% of all infighting episodes, followed by ANF (15%), ASIM (8%), and Jaish al-Islam (JAI) and the Kurdish YPG (7% each). FSA-affiliated factions were involved in about 6% of the infighting episodes. Our dyadic analysis is undirected, meaning we do not distinguish between who began the hostilities and who was merely defending.

STATE SPONSORSHIP at the node-level (rebel group) is simply coded as 1 if the group had a state sponsor at any point during the period 2013-2015. STATE SPONSORSHIP at the dyad level is coded as 1 if the two members of the dyad had any overlapping state sponsors. The major state sponsors of rebels have been Turkey, Saudi Arabia, Qatar, Jordan, and the United States (Phillips, 2016). Some groups have multiple sponsors. There are many organizations with no or unknown state sponsors. These include ISIL and ANF, but also many of the smaller groups.

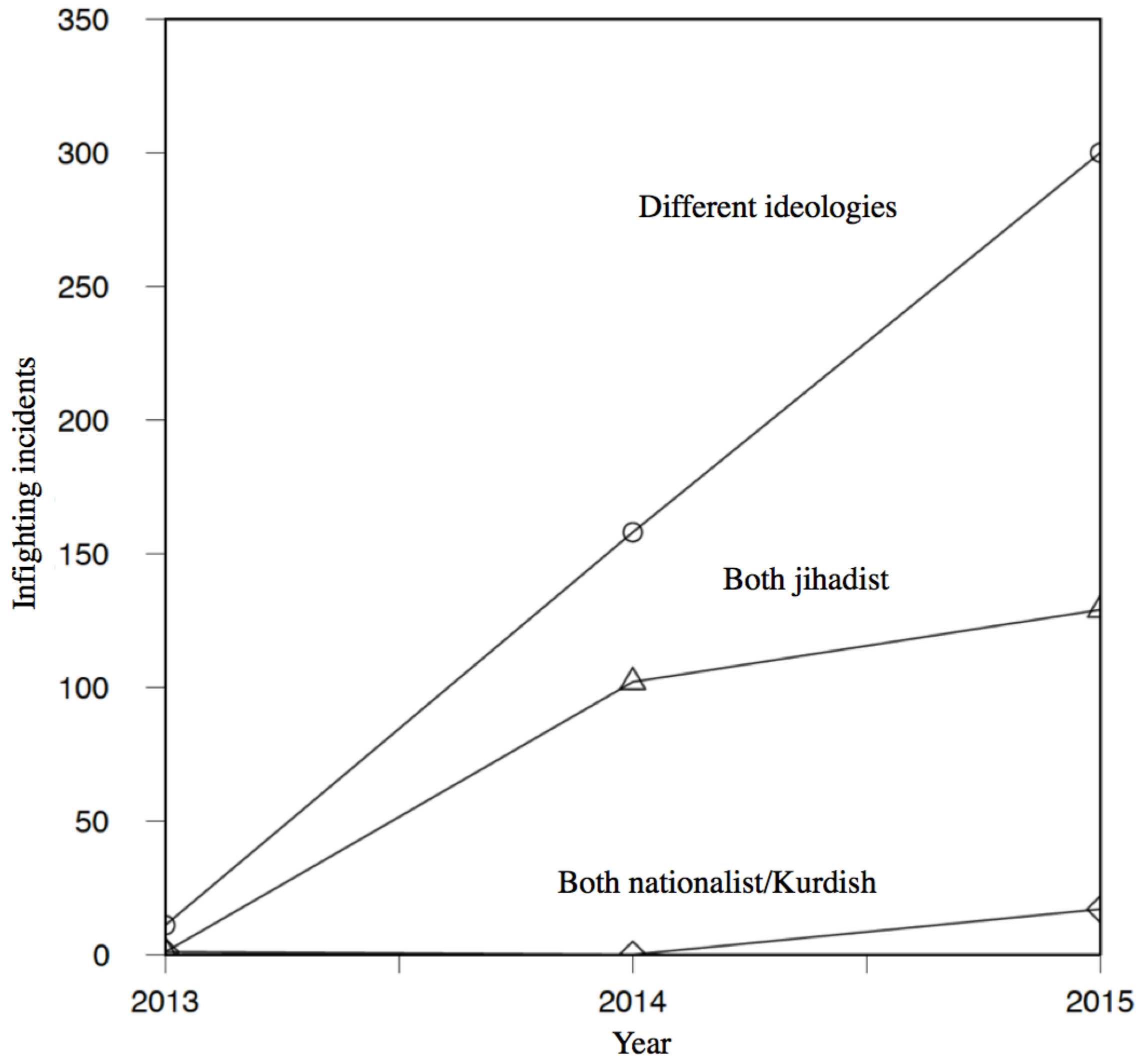

Figure 1. Infighting incidents by ideological dyads
Ideology is binarized: 3 or greater equals jihadist; lower than 3 equals nationalist/Kurdish factions.

We included two additional control variables for use in our network regression analysis. OPERATIONAL LOCATION may be an important practical factor in driving infighting. Rebels that are in close proximity to each other can more easily fight than those who are far apart. ISLAMIC STATE is a binary node level variable, coded one for ISIL and 0 if not. ISIL had the largest number of infighting ties in our network, and it is a highly ideological rebel group. We include this control to ensure ISIL was not the sole driver behind our ideology findings.

Network description

Figure 2 displays a visual representation of the network of infighting ties. Each line or tie denotes a single infighting episode. Observe that ISIL, along with ANF, ASIM, JAI, and YPG, form the center of the network, which means that they are the most frequent participants in interrebel conflicts. Table II in *Supplemental Material* displays the descriptive statistics of that network and our variables.

[Figure 2 about here]

Figure 2. The network of Syrian militant infighting ties
A tie indicates a single infighting relationship with another group.

Results

Assortativity and network simulation

A network whose elements correspond to the number of clashes between groups forms the basis of our analysis. Table I shows the observed assortativity values of the infighting network with respect to the ideology and power variables for the complete data time period (2013-2015). We discuss the ideology results first. Considering 2013-15, the observed assortativity for every ideology variable is more negative than its assortativity in the null simulation, indicating heterophily. All three ideology components and their average are found to have highly significant deviations from the null assortativity. They also have significant correlations with one or the other of the two eigenvectors with the most negative eigenvalues. This alignment with salient structural features in the network suggests that ideological heterophily plays an important role in shaping patterns of conflict. The 2014 and 2015 assortativity results show a similar pattern to 2013-15 with all the ideology variables again significant. For 2014, all of the components best correlate with the second most negative eigenvector. However, for 2015, TERRITORIAL ASPIRATION best correlates with the most negative eigenvector, a change which suggests that, as concerns over the disintegration of Syria grew, the salience of TERRITORIAL ASPIRATION intensified. Overall, we conclude that the high significance of the ideology assortativity for the full data and both individual years provides strong support for H2, that the likelihood of infighting increases with the ideological distance between groups.

[Table I about here]

Table I. Assortativity, community structure, and simulation results

	ASSORTATIVITY				EIGENVECTOR CORRELATION			VARIABLE SIMULATION	
VARIABLE	α	α_{null}	σ_{null}	*p*	*EV*	*r*	*p*	l_S/l	*CI*
2013-2015									
N = 30, *m* = 697									
Conflict Frame	-0.582(–)***	-0.265	0.036	<.0001	2MN	.454*	.012	2.5	(1.8,3.5)
Ideal Polity	-0.342(–)***	-0.148	0.035	<.0001	2MN	.436*	.02	--	--
Territorial Asp.	-0.572(–)***	-0.287	0.035	<.0001	1MN	.394*	.03	2.2	(0.7,3.3)
Average Ideol.	-0.616(–)***	-0.284	0.036	<.0001	2MN	.409*	.02	1.7	(1.2,2.3)
Power	-0.308(–)**	-0.209	0.035	.005	2MN	.238	.21	2600	(1000,4000)
2014									
N = 22, *m* = 260									
Conflict Frame	-0.529(–)***	-0.235	0.059	<.0001	2MN	.532*	.011	--	--
Ideal Polity	-0.269(–)**	-0.126	0.055	.006	2MN	.455*	.03	--	--
Territorial Asp.	-0.540(–)***	-0.302	0.056	<.0001	2MN	.470*	.03	3.6	(2.1,5.2)
Average Ideol.	-0.554(–)***	-0.264	0.058	<.0001	2MN	.545**	.009	3.3	(1.6,5.3)
Power	-0.263(+)	-0.298	0.056	.54	2MP	.321	.14	--	--
2015									
N = 24, *m* = 424									
Conflict Frame	-0.634(–)***	-0.290	0.045	<.0001	2MN	.442*	.03	1.8	(1.2,2.4)
Ideal Polity	-0.390(–)***	-0.167	0.044	<.0001	2MN	.434*	.03	--	--
Territorial Asp.	-0.610(–)***	-0.291	0.045	<.0001	1MN	.465*	.02	1.4	(0.8,2.2)
Average Ideol.	-0.671(–)***	-0.307	0.044	<.0001	1MN	.431*	.04	1.4	(0.9,1.9)
Power	-0.326(–)**	-0.185	0.046	.003	2MN	.360	.08	--	--

Variables are displayed under the corresponding time periods (*N* = no. groups, *m* = no. ties). For Assortativity: α is the assortativity of the observed network where the negative sign in parentheses indicates than α is less than α_{null} corresponding to heterophily (positive sign connotes homophily); α_{null} and σ_{null} are respectively the mean and standard deviation of the assortativity in the null simulation taken over 10,000 runs; the p-value *p* is the (two-tailed) fraction of runs exceeding $|\alpha - \alpha_{null}|$. For Eigenvector Correlation: *EV* is which one of the two most dominant eigenvectors has maximum correlation *r* with the variable (for heterophily, 1MN: most negative, 2MN: 2nd most neg.; for homophily, 2MP: 2nd most positive); *p* is the p-value of *r*. For Variable Simulation: l_S is the mean suppression length (heterophily) and l is the mean interaction length (homophily) at which the minimum error occurs and CI is the 95% confidence interval (blank entries signify the absence of a clear minimum); 1,000 runs at each point (ranging from 0.1 to 6 in 0.1 increments for ideology variables; from 500 to 25,000 in increments of 500 for power) were used to generate 1000 resamples of size 50 with replacement and then the l or l_S which minimized the squared error between the observed and simulated networks for each resample was found. *p < .05, **p < .01 ***p < .001.

As all ideology variables display heterophily, the suppression lengths, within which infighting is less likely, are reported for the ideology simulations with variable-based interactions. For the 2013-15 data period, well-defined suppression lengths for the heterophily simulation are found for TERRITORIAL ASPIRATION, CONFLICT FRAME, and AVERAGE IDEOLOGY. That $l_S = 1.7$ for Average Ideology indicates that the probability of infighting becomes substantially larger when the ideological distance between the groups in a dyad exceeds about half the full ideology scale. Note that IDEAL POLITY does not have a well-defined suppression length consistent with its relatively small magnitude assortativity. The 2015 time period is similar to the full dataset, but the suppression lengths are shorter indicating a narrowing of the ideological zone for which infighting is substantially less probable. For 2014, however, only TERRITORIAL ASPIRATION and AVERAGE IDEOLOGY exhibit well-defined suppression lengths although at relatively large values over 3. The difference between 2014 and 2015 parallels the decrease in assortativity to more negative values in the latter period.

Figure 3 visualizes the infighting network using the two least-ranked eigenvectors as node coordinates. The nodes are shaded with respect to their AVERAGE IDEOLOGY scores. The dominant pattern is represented by the most-negative eigenvector and shown on the vertical axis. It essentially corresponds to ISIL arrayed against everyone else. It is the second most negative eigenvector, shown on the horizontal axis, that best correlates with AVERAGE IDEOLOGY. This eigenvector indicates that AVERAGE IDEOLOGY tends to pit sectarian jihadists (ANF, ISIL) on the right against secular groups (YPG, Hazm) on the left. Powerful Salafist groups (ASIM, JAI) are found in the middle. Although the groups on the left show diversity in TERRITORIAL ASPIRATION, they are more uniform with respect to CONFLICT FRAME and IDEAL POLITY, being less sectarian and less Islamist than the jihadists on the right; an observation also consistent with those ideology variables correlating with the second most negative eigenvector while TERRITORIAL ASPIRATION best correlates with the most negative eigenvector. The left grouping is anchored by secular, non-sectarian groups (Syrian Revolutionaries Front and YPG) at the extreme whereas the highly sectarian ANF anchors the jihadist side.

[Figure 3 about here]

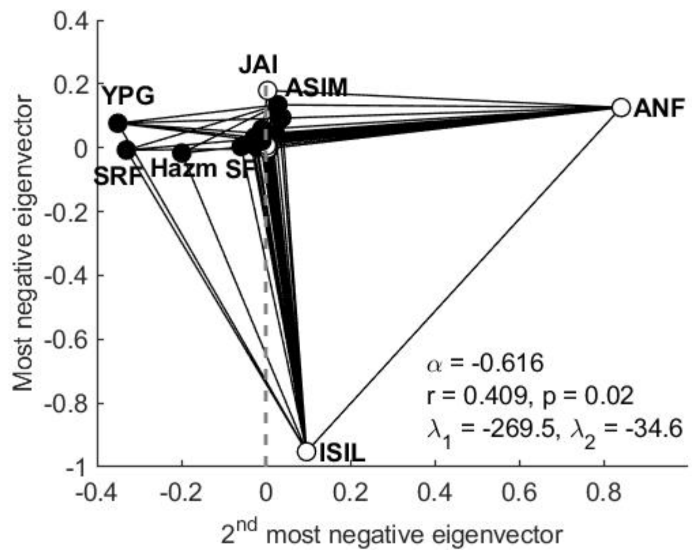

Figure 3. The infighting network structure for the 2013-2015 period

Solid circles indicate groups with AVERAGE IDEOLOGY < 3 (open circles ≥3). Links between groups indicate at least one clash. Assortativity (α) value of AVERAGE IDEOLOGY and correlation (*r*) and p-value with 2nd Most Negative Eigenvector shown. 1st and 2nd most negative eigenvalues denoted by λ_1 and λ_2. Vertical dashed line marks division into communities.

Turning to POWER, the difference between the observed POWER assortativity and the null mean value for 2013-15 is negative indicating heterophily and also highly significant thereby supporting the asymmetrical power hypothesis H1a that strength disparity tends to increase infighting. Although POWER does not correlate well with either of the two most negative eigenvectors, the variable-based simulation does yield a well-defined suppression length of 2,600. Heterophily with respect to power is also indicated by the significant negative assortativity deviation in 2015. The 2014 period, however, shows no significant effect of power and it is greater than the null assortativity indicating a tendency toward homophily rather than heterophily. Consequently, we conclude that there is strong evidence for the asymmetrical power dynamic although it is less consistent across analysis types and time than ideology.

Network regression results

Table II displays the results of network regression analysis. Hypotheses 1a and 1b are evaluated through examining a difference in POWER. A positive and statistically significant relationship with INFIGHTING would provide evidence in favor of H1a, power asymmetry: as the difference in size between two groups grows, they become more likely to fight. Conversely, a negative and statistically significant relationship would demonstrate support for H1b, power symmetry.

[Table II about here]

Table II. AME regression results

Beta values	*Model 1* Power	*Model 2* Ideology	*Model 3* Sponsor	*Model 4* Full model, all groups	*Model 5* Full model, Cont. for ISIL
Intercept	-0.054 (0.084)	-0.063 (0.082)	-0.01 (0.082)	-0.082 (0.298)	-0.200 (0.328)
Node-level variables					
Average ideology (1 to 5)				-0.003 (0.042)	-0.002 (0.005)
Power (0.5-40)				-0.005 (0.005)	-0.002 (0.005)
Sponsorship (Y/N)				-0.003 (0.116)	0.030 (0.125)
ISIL (Y/N)					1.449** (0.557)
Dyad-level variables					
Ideological difference (0-4)		0.035*** (0.010)		0.034*** (0.010)	0.033** (0.011)
Power difference (0-39.5)	0.004* (0.002)			0.005* (0.002)	0.005* (0.002)
Sponsor overlap (Y/N)			-0.03 (0.025)	-0.005 (0.025)	-0.009 (0.026)
Location overlap (Y/N)				0.067* (0.032)	0.071* (0.034)
Variance parameters					
Pmean (va)	0.044	0.042	0.042	0.067	0.050
Pmean (ve)	0.027	0.027	0.027	0.246	0.028
psd (va)	0.013	0.012	0.012	0.021	0.016
psd (ve)	0.001	0.001	0.001	0.012	0.002
N	30	30	30	30	30

The power range is in units of a thousand. DV is the square root of the count of infighting incidents. Statistical significance denoted at the 90%, 95%, 99% and 99.9% level: *p < .05, **p < .01 ***p < .001.

We see evidence for H1a in Table II: difference in size difference in size between groups generally increases the likelihood of INFIGHTING. We note that while this finding has the largest substantive impact, it does not achieve statistical significance in all models run: in binary outcome variables and a raw count, it does not achieve statistical significance at all (see *Supplemental Material*).

Hypothesis 2 is evaluated using the IDEOLOGICAL DISTANCE variable, defined as the difference between AVERAGE IDEOLOGY values of the groups in a dyad. Table II shows that IDEOLOGICAL DISTANCE displays a positive and statistically significant relationship regardless of which other variables at the dyad or node level are included, thereby indicating a greater tendency for infighting among ideologically dissimilar groups. This variable is by far the most consistent in terms of achieving statistical significance (regardless of model used, see *Supplemental Material)* and has a relatively large substantive impact (as compared with say location or sponsorship), lending strong support to H2. It even holds in terms of direction and significance of beta values in Model 5, which controls for ISIL at the node level. ISIL, the most frequent participator in infighting, has a strong effect on the network: this is manifest by the most negative eigenvector as shown in Figure 3, and by a large effect size and statistically significant result in Model 5. Thus, even when controlling for ISIL, ideology still has a statistically significant effect on the likelihood of infighting.

Hypothesis 3 is evaluated using shared STATE SPONSORSHIP. A negative and statistically significant relationship between STATE SPONSORSHIP and INFIGHTING would provide evidence in favor of H3 – that SHARED SPONSORSHIP makes INFIGHTING less likely. We find no evidence of a relationship between shared sponsorships at the dyad level and infighting (even when controlling for having or lacking a sponsor at the node level). The absence of a relationship could reflect the limited control that state sponsors have over their clients or a deliberate strategy to hedge by betting on multiple groups without concern over their collective cohesiveness. Rubin (2002: 198) noted this pattern in the case of Pakistan's support for the Afghan Mujahidin during their anti-Soviet jihad; and Staniland (2014: 163) in the case of India's support of Tamil factions in Sri Lanka during the 1980s.

ERGM results in the *Supplemental Material* confirm our findings concerning group size and ideology (though as with the AMEN analysis, power is less consistent across model specifications). Together, these three methods provide robust support for the ideological distance hypothesis, and support for power asymmetry shaping Syria's interrebel conflicts. We also find evidence that shared location makes infighting more likely.

Discussion and conclusions

We addressed the puzzle of interrebel wars at the dyadic level and tested it using data from the Syrian civil war. The results of our assortativity, community structure detection and network simulation, and AME regression indicate that two variables—ideological difference and power asymmetry—predict which rebel dyads are at most risk for infighting. Ideologically opposed groups have a higher propensity for infighting than ideologically proximate ones, and groups of disparate strength are more likely to fight with each other than ones of comparable power. These findings suggest that power and ideology need not be viewed as competing explanations of rebel infighting. Instead, they complement one another. For example, the power variable may explain *why* rebels engage in fratricide, including the quest for security and hegemony in a competitive landscape, and our analysis demonstrates that they are more likely to engage in war with groups of dissimilar size. Ideology, though, appears to tell us *with whom* groups are likely to fight in order to achieve their power aims. The greater the ideological distance between groups, the more likely they are to fight one another.

Our ideology findings may raise an endogeneity challenge. It could be asserted that infighting driven by power considerations compels groups to accentuate their ideological divides to justify or motivate the conflict. In this scenario, conflictual relationships drive the ideological distance exhibited in infighting networks, not the other way around. This objection assumes that militant groups arise as ideological blank slates, contrary to the fact that the founders of such groups often have strong ideological orientations from the outset. One of the key insights from Staniland's (2014: 33) social institutional theory of rebellion is that insurgents 'draw upon prewar political life in order to quickly form organizations that can handle the strains of violence.' Many of the individuals who would go on to form Syria's major Islamist rebel groups were actually in jail at the start of the revolution due to their prior Islamist activism and then subsequently released (Lister, 2015: 53-55).

Additionally, ideological manifestos and political programs, an important element of our coding, are typically issued by groups shortly after their formation. Their ideological statements, therefore, are biased toward a time before these groups have entered into interrebel wars. In other words, a variable biased toward the earlier period in the data cannot significantly correspond to patterns of conflict over two distinct years, 2014 and 2015, as well as the entire time period and yet be causally dependent on the infighting ties which are skewed toward a later period in the data (see Table 1). Table 1 also illustrates that ideology corresponded to the second most negative eigenvector, indicating that ideological differences are a greater force than power asymmetries in shaping the overall pattern of the network. Furthermore, ideology cannot be epiphenomenal to power since they are simultaneously significant in Models 4 and 5 of the AME regression analysis—i.e. when we control for power, ideology is still statistically significant and remains significant in all models. Nor can ideology simply be a function of overlapping state sponsorship since no significant relationship was found between the latter and the infighting network. Finally, ISIL is not driving the relationship as this finding exists even when controlling for this group in Model 5.

Another possible challenge to our results stems from our use of militant claims of infighting. It can be argued that militants preferentially reveal clashes with ideologically distant groups and conceal the ones with their ideological kin, and thus our data underrepresents infighting among ideological brothers. Of course, such self-censorship may occur, but to raise this objection beyond conjecture, one must estimate its frequency, a difficult task given the lack of ground truth. The use of militant claims enables construction of a dataset that is large enough to employ network analysis for a single conflict, an enterprise that can be pursued much less robustly using standard conflict datasets like UCDP. Although comparison with the UCDP data is difficult, when searching for missing infighting dyads (here, meaning pairs that fight at least once) among our ten most prominent groups (by degree), the UCDP data provides no evidence of a missing dyad. If the concealment of infighting between ideologically proximate groups was frequent, then one would expect that at least one such pair would be found in UCDP that is absent from our data. That there is none leaves little evidentiary basis to support the self-censorship objection. Furthermore, our data reflects many fighting episodes between ISIL and ANF, two ideologically similar movements. This suggests that the presumption that rebels mask their infighting when it is politically inconvenient is contradicted repeatedly in at least one prominent conflict dyad.

Lastly, one might object that our analysis leaves out the state, an important factor in shaping interrebel conflicts. The state's accommodative arrangements can give some rebel groups a greater degree of freedom to attack their rivals because they are less concerned about fighting the state. For example, the Syrian regime has been accused of deliberately neglecting ISIL in its targeting policy so that a tacit alliance between the state and ISIL allowed the latter to concentrate its fighting resources on wiping out its rivals. Although accounting for selective targeting of rebel groups by the state would add to the substantive analysis of infighting, doing so is empirically difficult as the Syrian regime typically made claims of attacking "terrorists" rather than specific groups.

We assess the impact of omitting the state through consideration of ISIL's effect on the network. Given that ISIL is responsible for 41% of all the infighting ties, the potential collusion between the Assad regime and ISIL would be the most consequential confounding factor arising from the omission of the state. However, Figure 3 shows that the second most-negative eigenvector best correlates with AVERAGE IDEOLOGY in the full data period and not the most-negative eigenvector, corresponding to ISIL arrayed against all the other groups – a result suggesting that ISIL is not the dominant driver of our ideological distance finding. Furthermore, removing ISIL from the network yields the best correlation with the most-negative eigenvector ($p = .02$), and the assortativity test still strongly supports ideological heterophily ($p < .0001$). That ideological heterophily persists in the pattern of infighting after removing ISIL provides strong evidence that infighting cannot be attributed to selective targeting by the state or its collusion with particular groups.

Our empirical findings in the Syrian case add weight to a burgeoning body of scholarship that makes the case that 'ideological considerations play a prominent role in guiding insurgent decision making' (Hirose, Imai & Lyall, 2017: 48). This literature has empirically demonstrated how ideological variables can explain important conflict processes in civil wars, such as anti-civilian atrocities (Straus, 2015) and mobilization effectiveness (Ugarriza & Craig, 2012; Costalli & Ruggeri, 2015). Our article shows that ideology also matters to rebel infighting and gives added credence to those who call for bringing armed politics back into the study of civil conflicts (Staniland, 2015; Balcells, 2017). Allowing for both power and ideology to impact insurgent factional dynamics should improve our understanding and anticipation of conflict trajectories in fragmented civil wars.

Replication data

The dataset, computational scripts, and Supplemental Material can be found at http://www.prio.org/jpr/datasets.

Acknowledgments

We would also like to thank Naazneen Barma, Cassy Dorff, Peter Krause and Zane Kelly.

Funding

This research was funded by the Office of Naval Research under grants N00014-15-1-2549 and N00014-16-1-2919.

References

Akcinaroglu, Seden (2012) Rebel interdependencies and civil war outcomes. *Journal of Conflict Resolution* 56(5): 879-903.

Bakke, Kristin; Kathleen Gallagher Cunningham & Lee JM Seymour (2012) A plague of initials: Fragmentation, cohesion, and infighting in civil wars. *Perspectives on Politics* 10(2): 265-283.

Balcells, Laia (2017) *Rivalry and Revenge: The Politics of Violence during Civil War*. New York: Cambridge University Press.

Bapat, Navin & Kanisha D Bond (2012) Alliances between militant groups. *British Journal of Political Science* 42(4): 793-824.

Christia, Fotini (2012) *Alliance Formation in Civil Wars*. New York: Cambridge University Press.

Costalli, Stefano & Andrea Ruggeri (2015) Indignation, ideologies, and armed mobilization: Civil war in Italy, 1943-45. *International Security* 40(2): 119-157.

Cranmer, Skyler; Philip Leifeld, Scott D McClurg & Meredith Rolfe (2017) Navigating the range of statistical tools for inferential network analysis. *American Journal of Political Science* 61(1): 237-251.

Cunningham, Kathleen Gallagher; Kristin M Bakke & Lee JM Seymour (2012) Shirts today, skins tomorrow: Dual contests and the effects of fragmentation in self-determination disputes. *Journal of Conflict Resolution* 56(1): 67-93.

Desmarais, Bruce A & Skyler J Cranmer (2017) Statistical inference in political networks research. In: Jennifer Nicoll Victor, Alexander H Montgomery & Mark Lubell (eds) *The Oxford Handbook of Political Networks*. New York: Oxford University Press, 203-220.

Dorussen, Han; Erik A Gartzke & Oliver Westerwinter (2016) Networked international politics: Complex interdependence and the diffusion of conflict and peace. *Journal of Peace Research* 53(3): 283-291.

Fjelde, Hanne & Desirée Nilsson (2012) Rebels against Rebels: Explaining violence between rebel groups. *Journal of Conflict Resolution* 56(4): 604-628.

Griffiths, Ryan D (2016) *Age of Secession: The International and Domestic Determinants of State Birth*. New York: Cambridge University Press.

Gutiérrez Sanín, Francisco & Elisabeth Jean Wood (2014) Ideology in civil war: Instrumental adoption and beyond. *Journal of Peace Research* 51(2): 213-226.

Hafner-Burton, Emilie M; Miles Kahler & Alexander H Montgomery (2009) Network analysis for international relations. *International Organization* 63(3): 559-592.

Hirose, Kentaro; Kosuke Imai & Jason Lyall (2017) Can civilian attitudes predict insurgent violence? Ideology and insurgent tactical choice in civil war. *Journal of Peace Research* 54(1): 47-63.

Hoff, Peter D (2015) Dyadic data analysis with amen (https://arxiv.org/abs/1506.08237).

Hoover Green, Amelia (2016) The commander's dilemma: Creating and controlling armed group violence. *Journal of Peace Research* 53(5): 619-632.

International Crisis Group (2012) Tentative Jihad: Syria's fundamentalist opposition (https://d2071andvip0wj.cloudfront.net/131-tentative-jihad-syria-s-fundamentalist-opposition.pdf).

International Crisis Group (2013) Syria's metastasising conflicts (https://www.crisisgroup.org/middle-east-north-africa/eastern-mediterranean/syria/syria-s-metastasising-conflicts).

International Crisis Group (2014) Flight of Icarus? The PYD's precarious rise in Syria (https://www.crisisgroup.org/middle-east-north-africa/eastern-mediterranean/syria/flight-icarus-pyd-s-precarious-rise-syria).

Jones, Seth (2017) *Waging Insurgent Warfare: Lessons from the Vietcong to the Islamic State*. New York: Oxford University Press.

Krause, Peter (2013/14) The structure of success: How the internal distribution of power drives armed group behavior and national movement effectiveness. *International Security* 38(3): 72-116.

Krause, Peter (2017) *Rebel Power: Why National Movements Compete, Fight, and Win*. Ithaca, NY: Cornell University Press.

Kydd, Andrew & Barbara F Walter (2002) Sabotaging the peace: The politics of extremist violence. *International Organization* 56(2): 263-296.

Lichbach, Mark Irving (1995) *The Rebel's Dilemma*. Ann Arbor, MI: University of Michigan Press.

Lister, Charles R (2015) *The Syrian Jihad: Al-Qaeda, the Islamic State, and the Evolution of an Insurgency*. New York: Oxford University Press.

Maoz, Zeev (2010) *Networks of Nations: The Evolution, Structure, and Impact of International Networks, 1816–2001*. New York: Cambridge University Press.

McLauchlin, Theodore & Wendy Pearlman (2012) Out-group conflict, in-group unity? Exploring the effect of repression on intramovement cooperation. *Journal of Conflict Resolution* 56(1): 41-66.

Metternich, Nils; Cassy Dorff, Max Gallop, Simon Weschle & Michael D Ward (2013) Antigovernment networks in civil conflicts: How network structures affect conflictual behavior. *American Journal of Political Science* 57(4): 892-911.

Minhas, Shahryar; Peter D Hoff & Michael D Ward (2016) A new approach to analyzing coevolving longitudinal networks in international relations. *Journal of Peace Research* 53(3) 491-505.

Newman, Mark EJ (2006) Finding community structure in networks using the eigenvectors of matrices. *Physical Review E* 74(3): 1-22.

Oppenheim, Benjamin; Abbey Steele, Juan Vargas & Michael Weintraub (2015) True believers, deserters, and traitors: Who leaves insurgent groups and why. *Journal of Conflict Resolution* 59(5): 794-823.

Phillips, Christopher (2016) *The Battle for Syria: International Rivalry in the New Middle East*. New Haven, CT: Yale University Press.

Pischedda, Costantino (2018) Wars within wars: How windows of opportunity and vulnerability cause inter-rebel fighting. *International Security* 43(1): 138-176.

Rubin, Barnett R (2002) *The Fragmentation of Afghanistan*. New Haven, CT: Yale University Press.

Salehyan, Idean (2010) The delegation of war to rebel organizations. *Journal of Conflict Resolution* 54(3): 493-515.

Salehyan, Idean; David Siroky & Reed M Wood (2014) External rebel sponsorship and civilian abuse: A principal-agent analysis of wartime atrocities. *International Organization* 68(3): 633-661.

Seymour, Lee (2014) Why factions switch sides in civil wars: Rivalry, patronage, and realignment in Sudan. *International Security* 39(2): 92-131.

Seymour, Lee; Kristin M Bakke & Kathleen Gallagher Cunningham (2016) E pluribus unum, ex uno plures: Competition, violence, and fragmentation in ethnopolitical movements. *Journal of Peace Research* 53(1): 3-18.

Staniland, Paul (2014) *Networks of Rebellion: Explaining Insurgent Cohesion and Collapse*. Ithaca, NY: Cornell University Press.

Staniland, Paul (2015) Militias, ideology, and the state. *Journal of Conflict Resolution* 59(5): 770-793.

Straus, Scott (2015) *Making and Unmaking Nations: War, Leadership and Genocide in Modern Africa*. Ithaca, NY: Cornell University Press.

Tamm, Henning (2016) Rebel leaders, internal rivals, and external resources: How state sponsors affect insurgent cohesion. *International Studies Quarterly* 60(4): 599–610.

Ugarriza, Juan & Matthew J Craig (2012) The relevance of ideology to contemporary armed conflicts: A quantitative analysis of former combatants in Colombia. *Journal of Conflict Resolution* 57(3): 445-477.

Ward, Michael; Katherine Stovel & Audrey Sacks (2011) Network analysis and political science. *Annual Review of Political Science* 14: 245-264.

Zech, Steven & Michael Gabbay (2016) Social network analysis in the study of terrorism and insurgency: From organization to politics. *International Studies Review* 18(2): 214-243.

EMILY KALAH GADE, b. 1985, PhD in Political Science (University of Washington, 2017); Acting Assistant Professor of Political Science and Moore/Sloan Innovations in Data Science Postdoctoral Fellow, University of Washington (2017 -); research area: political violence, data science.

MOHAMMED M HAFEZ, b. 1970, PhD in International Relations (London School of Economics, 2000); Associate Professor of National Security Affairs, Naval Postgraduate School (2008 -); research area: Islamist political violence.

MICHAEL GABBAY, b. 1963, PhD in Physics (University of Chicago, 1997); Senior Principal Physicist, Applied Physics Laboratory, University of Washington (2009 -); research area: political networks.

Appendix

We describe the formalism used to calculate the assortativity, relate it to the eigenvector spectrum of the modularity matrix, and formalize the network simulation. We consider a symmetric network with N nodes represented by the adjacency matrix with components A_{ij}, equal to the number of ties between nodes i and j ($A_{ij} = A_{ji}$). The degree of node i is the sum of its ties, $k_i = \sum_{j=1}^{N} A_{ij}$. The total number of ties in the network is m. These network quantities are used to define the components of the modularity matrix **B**,

$$B_{ij} = A_{ij} - \frac{k_i k_j}{2m}, \qquad (1)$$

which is the difference between the observed tie strength and what would be expected from a null process in which ties are formed in proportion to the product of node degrees without regard to any interactions driven by node variables. Assortativity is the standard measure used to assess homophily of network tie formation with respect to a scalar variable (as we operationalize power and ideology).

The assortativity a that the network displays with respect to a node variable x can be related to the modularity matrix via

$$\alpha = \frac{\sum_{i,j} B_{ij} x_i x_j}{\sum_{i,j} A_{ij} x_i^2 - (k_i k_j / 2m) x_i x_j} , \qquad (2)$$

which is equivalent to the correlation of x over ties (Newman, 2006).

The spectrum of the modularity matrix is defined via $\mathbf{B}\mathbf{u}_\nu = \lambda_\nu \mathbf{u}_\nu$ where λ_ν is the eigenvalue corresponding to eigenvector $\mathbf{u}_\nu$ with eigenvectors indexed in order of decreasing eigenvalue. Newman (2006) shows that the assortativity can be expanded as a sum of the modularity matrix eigenvalues where the weight associated with each λ_ν is proportional to square of the inner product of the vector formed by the x_i and the associated eigenvector $\mathbf{u}_\nu$, $(\sum_{i=1}^{N} u_i^{(\nu)} x_i)^2$. Homophily, therefore, can be manifested by significant correlations of the variable with a highly ranked eigenvector (of positive eigenvalue), not just the leading one. Heterophily, on the other hand, is manifested by significant correlations with low-ranked eigenvectors, i.e., those with the most negative eigenvalues.

To formulate the null and variable-based simulations, we denote the maximum possible degree of node i by D_i (its degree in the empirical network) and its degree at iteration step n by $k_i(n)$. The deficit between the maximum and current degrees is then $D_i - k_i(n)$. Accordingly, the probability of tie formation between nodes i and j at step n is

$$p_{ij}(n) = \begin{cases} K_n(D_i - k_i(n))(D_j - k_j(n))\, f(x_i - x_j), & i < j \\ 0, & i \geq j. \end{cases} \qquad (3)$$

K_n is a normalization constant so that $\sum_{i,j} p_{ij}(n) = 1$. The bottom line prevents self-ties (the $i > j$ case need not be considered separately because the network is symmetric). The function $f(x_i - x_j)$ accounts for the dependence of the probability upon the difference of the variable x between the two nodes. We make three choices for $f(x_i - x_j)$ corresponding to: (1) the null simulation, $f(x_i - x_j) = 1$; (2) the homophily simulation, $f(x_i - x_j) = \exp\left(-0.5(x_i - x_j)^2 / l^2\right)$ (a Gaussian); and (3) the heterophily simulation, $f(x_i - x_j) = 2 - \exp\left(-0.5(x_i - x_j)^2 / l_S{}^2\right)$.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*

2. REPORT TYPE
Final Technical

3. DATES COVERED *(From - To)*
5/1/2015 - 9/30/2018

4. TITLE AND SUBTITLE
Critical Transitions and Adaptation in Group Dynamics

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-15-1-2549

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Michael Gabbay

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Washington – Applied Physics Laboratory
4333 Brooklyn Avenue NE
Seattle, WA 98105-6613

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
ONR EXP WARFARE & COMBATING TERRORISM DEPT (Code 30)
875 North Randolph Street
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT:

Distribution Statement A: Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This report describes research exploring group opinion dynamics and aspects of militant group decision making. A novel theory of how group discussion leads to extreme decisions that incorporates research on decision making under risk and uncertainty was developed. In this theory, the focal frame of discussion, such as the subjective probability of an outcome, can systematically facilitate majority formation toward the extreme. Relatedly, a novel model of opinion network dynamics was developed which enables the emergence of proximate majorities and distinguishes between opinion and rhetoric. The model is in quantitative agreement with the results of a recent group experiment. Simulations also display a sharp transition between convergence to a common opinion and divergence into opposed camps. A theoretical framework for the role of ideology in militant decision making was developed and tested using data on insurgent groups in Syria. Using data collected on militant groups in Northern Ireland, factors responsible for escalation or de-escalation decisions were identified.

15. SUBJECT TERMS
Group decision making, opinion networks, social networks, terrorism, insurgency, modeling and simulation

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified	Unclassified	Unclassified	UU	147	Michael Gabbay
					19b. TELEPHONE NUMBER *(include area code)* (206) 543-1300